

**The nature of host plant recruitment by the sensory repertoire of *Sinorhizobium meliloti***

**Keith Karl Compton**

Dissertation submitted to the faculty of the Virginia Polytechnic Institute and State University in  
partial fulfillment of the requirements for the degree of

Doctor of Philosophy

in

Biological Sciences

Birgit E. Scharf, Chair

John Jelesko

Florian D. Schubot

Dorothea Tholl

August 10<sup>th</sup> 2020

Blacksburg, Virginia

Keywords: legume, chemotaxis, attractant, methyl accepting chemotaxis protein, symbiosis

CC BY-ND © 2020 Keith Karl Compton

# **The nature of host plant recruitment by the sensory repertoire of *Sinorhizobium meliloti***

**Keith Karl Compton**

## **ABSTRACT**

*Sinorhizobium meliloti* (*Ensifer meliloti*) is a bacterium that will exist saprotrophically in the soil and rhizosphere or as a differentiated bacteroid inside root nodules of the legume genera *Medicago*, *Melilotus*, and *Trigonella*. It exists in symbiosis when inside a host plant and will fix gaseous N<sub>2</sub> into ammonium for the plant. In return, a population of the bacteria is harbored inside the plant where it can proliferate beyond what would be possible in the rhizosphere or bulk soil. This symbiosis is a defining feature of the Fabaceae (legume) family, a clade that diverged approximately 60 million years ago and is now the 5<sup>th</sup> largest plant family by species count. Each legume species pairs with one or several strains of bacteria, referred to broadly as rhizobia. The rhizobia identify their proper host plant by a cocktail of secondary metabolites called flavonoids released from specific parts of the roots. Initiation of the symbiosis may only occur at the tips of young root hairs. Therefore, the means rhizobia take to localize themselves to these sites must be the inceptive step in the symbiotic interaction. The studies here examine the mechanisms and priorities rhizobia use to achieve this goal. Movement of bacteria is referred to as motility and is achieved via (in rhizobia, multiple) rotating flagella, proteinaceous extracellular appendages that propel the cell through liquid environments. On their own, flagella may only move but not guide the cell. Navigation is achieved through sensors that detect chemical attractant or repellent cues in the environment and an intracellular signaling system that relays information to appropriately control locomotion. This sensing is called chemotaxis. A research focus is directed on the sensing aspect of chemotaxis to understand which chemical compounds are the preferred attractants for *S. meliloti*. An emphasis is placed on those compounds released from germinating host seeds.

Chapter 2 spearheads our research goals by examining the chemotactic potential of host-derived flavonoids, the compounds that induce the symbiotic signaling in the rhizobial symbiont. While a logical place to start, this study reveals that our strain of rhizobia is not attracted to flavonoids. We determined that the best chemoattractants are hydrophilic in nature and that hydrophobic compounds, such as flavonoids, are not effective chemoattractants. In addition, we discuss the nature of chemotactic

agents and symbiosis inducers to fortify our understanding of how classes of compounds contribute to the rhizobia-plant interaction.

In chapter 3, we characterize the sensor protein, McpV, and its ligand profile for carboxylates. The protein is first screened using a high-throughput assay to test numerous possible ligands simultaneously. We confirm positive reactions using direct binding studies and quantify dissociation constants. Then, the phenotypic response to these ligands is measured using capillary chemotaxis assays, and the role *mcpV* plays in this response is confirmed using deletion mutants. Last, the symbiotic context is addressed by quantifying these ligands in exudates of the host alfalfa. These experiments show that McpV is a chemotactic sensor dedicated to detecting 2 – 4 C monocarboxylates. Only one of the compounds found in the ligand profile, glycolate, was detected in seed exudates, so the contribution of McpV to host sensing is yet to be expounded.

Chapter 4 follows the model of chapter 2 but is complicated when the ligand screen used previously gives ambiguous results. Using direct binding studies, we were able to confirm the true ligand amidst numerous false positives. Analytical gel filtration suggests that McpT exists as a dimer regardless of ligand binding. Capillary chemotaxis assays quantified the responses mediated by McpT to di- and tri-carboxylates, which were slightly weaker, but still on-par with the responses to McpV ligands. Strains with *mcpT* deletions showed strongly reduced, but in some cases, not abolished, chemotaxis to carboxylates.

Chapter 5 examines McpX – the chemoreceptor already known to be a sensor of quaternary ammonium compounds. This is a structural investigation into the binding of McpX to its ligands. A crystal structure of the ligand binding region of the protein is resolved to understand how ligands fit into the binding pocket of McpX and what determines its structurally diverse ligand profile. The contribution of certain residues to ligand binding are further probed using direct binding studies on single point variants of McpX.

The analysis of chemoreceptor functions hint at what kinds of molecules are most important to bacterial survival and reproduction. Knowing what the bacterium is tuned to seek out grants understanding of what niches they prefer, and how they thrive in those niches. For *S. meliloti* and other rhizobia, the preeminent niche is one in symbiosis with a host plant. The sum of this knowledge we have accrued with *S. meliloti* lends itself to agricultural goals of soil enrichment, legume inoculation, nutrient cycling, and environmentally safe and efficient crop fertilization.

## **The nature of host plant recruitment by the sensory repertoire of *Sinorhizobium meliloti***

**Keith Karl Compton**

### GENERAL AUDIENCE ABSTRACT

*Sinorhizobium meliloti* and other soil-dwelling bacteria termed rhizobia are crucial to the cultivation of leguminous crops such as alfalfa, soy, pea, lentil, peanut, and many more. The bacterium can be internalized by the plant host's roots where it will supply the plant with nitrogen. This is a great boon to crops when they need to accumulate more protein in seed stores, or for plants that survive in nutrient depleted soils. The bacterium must begin seeking out the host plant by sensing chemical cues. It can navigate to the proper location by using a process called chemotaxis. This process is centered around chemoreceptors that can be likened to the nose of the bacterium. Using these chemoreceptors, the bacterium will seek out compounds that benefits it – these are usually food sources. Identifying what each individual chemoreceptor senses allows us to understand what the bacterium needs to seek out to survive. We correlate this information with compounds that the plant secretes and find that many chemoreceptors have evolved to sense signals that will lead the bacterium to a plant root. This interaction is a key part of how the symbiosis is propagated and ultimately benefits the agriculture of leguminous plant

## Acknowledgement

This document contains the sum of my material efforts in the past few years. It may not be apparent, but I have been propped up, pushed, and encouraged the whole way. I confer my sincerest gratitude to everyone here. A part of this work is credited to each of you.

Dr. Scharf, there is no one at the university more dedicated to mentorship than you. Raising your students into exemplar scientists is your primary goal and it shows. You pushed me when I was strong and pulled me when I faltered. Few other PIs make their work as much about the pupil and I am deeply appreciative of that. I will treasure the skills and knowledge I cultivated with you for the rest of my life.

The members of the Scharf lab, you guys and gals made this place much more than a data factory. The advice and camaraderie have been a boon over the last seven years. A special thanks goes to everyone who lent an ear to my ramblings, incoherent half-thoughts, and gripes.

Benjamin Webb, a very special thanks goes to you. Over the past few years, I have needed to revisit your work and lab books. Reading your experimental notes as well as your personal thoughts and feelings has given me such an appreciation for the work you indirectly did for me. If not for the rabbit holes, pilot experiments, what-ifs, fool's errands, and shots-in-the-dark, this thesis would have been delayed considerably longer and with much more trouble. Thank you for all the things you did to establish the lab, the environment you cultured, and the tireless mentorship and unequivocal friendship. I am so pleased that you are building your own trail through life.

Mom and Dad, thank you for always trying your hardest to convince me that I am the smartest, handsomest, strongest, fastest, all around best boy in the world – even if all the evidence is to the contrary. My life is truly enriched because of the care you gave for me. All the work you put in to raising me could not have offered a better outcome. All my efforts and accomplishments in life are in part attributed to you. I love you both so much.

Claire, thanks for being the “cool adult” in my life. The holidays and stays with you, Joe, and Gooby mean the world to me. The Lellow Genie is here for you.

Julia, weren't we quite a pair? You're the reason I didn't grow up a complete luddite. Thanks for being my tour guide to the early internet. There are no better hands for Beau to be in. I am putting it here in writing that I am going to visit more often.

Grandpa Gene and Grandma Rose, like my parents you give my traits and talents an enormously high appraisal, but I know it is still out of love. I am blessed to have you both in my life.

Grandma Lorene, I am so happy you have found a good place with my parents and Val. I know you will all help each other out in your own way. I love you dearly.

Bunny, few days pass without my wistful thoughts turning to you. You were the embodiment of sweetness for me and I will always love you. Though you are gone, I still hold the peace from your cherished life.

Earl, you were a window into myself. You gave me the confidence to believe in my self-awareness and my understanding of who I am. I am glad we met. You also do some of the best cooking ever.

The Translational Plant Sciences community, thank you for building a system of support for me and providing me with so many opportunities. Ours is a special group at VT.

My friends in microbiology, both present and past, you all have been wonderful companions over the past few years, even though the topic of conversation usually turned morose. It was always fun.

To my thesis committee; John Jelesko, Florian Schubot, and Dorothea Tholl, thank you for the guidance and insight over the years. I know now that I had a stellar group of scientists to help me navigate these projects. Thank you all for always being on my side.

## Table of Contents

Chapter 1 – Introduction .....	1
REFERENCES .....	11
Chapter 2 - An updated perspective on <i>Sinorhizobium meliloti</i> chemotaxis to alfalfa flavonoids .....	21
ABSTRACT .....	22
INTRODUCTION .....	23
MATERIALS AND METHODS .....	25
RESULTS .....	28
DISCUSSION .....	31
REFERENCES .....	36
Chapter 3 - <i>Sinorhizobium meliloti</i> chemoreceptor McpV senses short chain carboxylates via direct binding .....	53
ABSTRACT .....	54
IMPORTANCE .....	54
INTRODUCTION .....	55
RESULTS .....	56
DISCUSSION .....	61
MATERIALS AND METHODS .....	64
ACKNOWLEDGEMENTS .....	68
REFERENCES .....	72
Chapter 4 - McpT is a broad range carboxylate chemoreceptor in <i>Sinorhizobium meliloti</i> .....	88
ABSTRACT .....	89
INTRODUCTION .....	90
RESULTS .....	91
DISCUSSION .....	96
MATERIAL AND METHODS .....	98
REFERENCES .....	104
Chapter 5 - Structure of the sensory domain of McpX from <i>Sinorhizobium meliloti</i> , the first known bacterial chemotactic sensor for quaternary ammonium compounds .....	124
ABSTRACT .....	125
INTRODUCTION .....	126
MATERIALS AND METHODS .....	128
RESULTS .....	131
DISCUSSION .....	135
REFERENCES .....	139

Chapter 6 – Final Discussion .....	152
REFERENCES .....	159



## List of Abbreviations

DSF: Differential Scanning Fluorimetry

ITC: Isothermal Titration Calorimetry

$K_d$ : Dissociation Constant

LBD: Ligand Binding Domain

MBP: Maltose-Binding Protein

MCP: Methyl-accepting Chemotaxis Protein

NF: Nod-Factor

PGPR: Plant-Growth Promoting Rhizobacteria

PR: Periplasmic Region

## Chapter 1 – Introduction

A common concept thrown around in biology classrooms is that microbes, especially bacteria, are ubiquitous. There are few better examples of this than the soil. Along with a cosmopolitan roster of fungi, nematodes, insects, and protists; bacteria daub every surface, and pervade every pore. When a plant puts down roots, it wades through nearly as many microbes as it does soil particles. Like animals, plants interact with these neighbors of circumstance in a number of sometimes surprising ways. The more infamous microbes antagonize the plant, causing diseases like galls from *Agrobacterium tumefaciens*, or wilting from *Erwinia tracheiphilia*. On the opposite end of the spectrum are organisms that benefit the plant by providing mineral nutrients. While a plant specializes in harvesting light to generate energy stores of fixed carbon, the soil microbes are best at scouring crevices and pockets in the soil to maximize the retrieval of mineral nutrients like potassium, nitrogen, calcium, and phosphorus. A trade is struck between the two sides. Some microorganisms take the practice to extremes in which they can only derive certain nutrients from a host, and thus are dependent on it for survival. Perhaps the best example of highly specific interactions is the rhizobia, a clade of soil bacteria that are capable of independent survival in the soil but also have the ability to be internalized by a specific host plant and become entirely dependent on the plant for persistence. The benefit to the plant comes when the bacteria transform into organelle-like structures that turn the normally inert  $N_2$  gas into  $NH_3$  – ammonia fertilizer. The legumes; pea, soy, clover, locust, acacia, vetch, lentil, and others, are the primary employers of this arrangement. This dissertation falls within the scope of the interplay between alfalfa (*Medicago sativa*), and *Sinorhizobium meliloti*. In particular, we ask how the bacterium might use host-sourced chemical cues to navigate to the proper location on the roots. A special interest will be targeted to the molecular mechanisms of chemical sensing.

### **Diversity of legumes**

The Fabaceae family is thought to have diverged approximately 60 million years ago (1). Nodulation with rhizobia is extant in nearly all clades within this family, as well as the closely related non-legume *Parasponia*. Nodulation is an involved process with many steps; at each step, legumes have evolved highly divergent variations that sum to a staggering array of diversity (2). One such example is the mechanism by which the bacteria initially enter, or infect, their host's

roots (Figure 1.1). The ancestral method is thought to be crack infection, whereby bacteria enter through cracks in the epidermis that naturally arise when lateral roots emerge. This mechanism of entry is shared with many pathogens. Other methods include intercellular invasion (where the bacteria penetrate the space in between cells to migrate to the root cortex) and infection threads (in which the plant forms a channel dedicated to transporting the bacteria into the cortical root layers). In any of these cases, the plant host is the determinant for entry, signaling, and sensing methods (3, 4). The onus of compliance is on the bacterium. Another important point of variation is the form and nature of the nodule. Although there are many outliers, nodules can be classified into two types by the presence of a persistent meristem (Figure 1.2). The determinate nodules lack a meristem, giving them a round, spherical appearance. Upon senescence, a determinate nodule will be entirely sloughed off and the incorporated bacteria are released from subjection. The indeterminate nodules maintain a meristem which creates an elongated, cylindrical shape that can often branch into a new nodule body. Again, and as with mechanisms of entry, the species of plant host dictates one or the other nodule type (5). The bacteria will merely fill the vessel. A plant may be promiscuous in the range of rhizobial species it will nodulate with. *S. meliloti* has been noted to nodulate *Melilotus* and *Trigonella* but our models and experiments are only concerned with its interaction with alfalfa.

### **Diversity of rhizobia**

Where legumes show diversity in their morphology, the rhizobia have a preponderance of genetic and taxonomic diversity. Traditionally, rhizobia broadly fit into the Rhizobiales order under the alpha-Proteobacteria. It should be noted here that nodulating bacteria form a paraphyletic group. Certain species of *Bradyrhizobium* and *Ensifer* do not appear to nodulate plants. Furthermore, many species related to the pathogenic *Burkholderia* under the beta-Proteobacteria class have been found to nodulate legumes. This is an emerging and fascinating side of symbiosis but is yet out of our scope (6). Rhizobia have as much or more breadth in their host ranges as the legumes do symbionts. *Rhizobium* sp. NGR234 for example was found to nodulate over 100 genera of legumes (7). The suite of genes that determine host compatibility are typically borne on large mega plasmids or chromids. These genes may be swapped to change host ranges between similar strains or species of rhizobia. Exchanges of symbiotic genes between strains that have highly different hosts results in bacteria that can partially infect a new host but can never achieve nitrogen fixation in a fully

mature nodule. Because the molecular determinants are a handful of transmissible genes, the same species of rhizobia can have different host ranges (4, 8-10). For this reason, strains, biovars, and recently, symbiovars are used to distinguish taxa for the purposes of symbiosis (11). Even at the genus level, rhizobia (and many bacteria in general) can be difficult to classify. The model organism *S. meliloti* was first described as *Rhizobium meliloti* until 2008. It was erroneously reassigned to a new genus, *Sinorhizobium*, until it was determined that this group was closely related to a preexisting genus, *Ensifer*, which has priority. Therefore, the proper name of this organism is *Ensifer meliloti* (12). In spite of this, the community of rhizobia and symbiosis researchers continues to use *Sinorhizobium meliloti* as a synonym. For the sake of conformity and simplicity, we have also adopted this name.

### ***S. meliloti* and its host, alfalfa**

*S. meliloti* is a gram-negative, aerobic to microaerophilic rod-shaped bacterium that is ubiquitous in the soil. Its genome is tripartite, composed of a 3.65 Mb chromosome, the 1.35 Mb pSymA megaplasmid, and the 1.68 Mb pSymB megaplasmid. While pSymA contains the genes for symbiosis and nitrogen fixation, pSymB harbors other genes, such as essential tRNAs that make the megaplasmid necessary for survival (13). It owes its motility to a peritrichous arrangement of flagella. In the absence of a plant host, the bacteria will survive on organic matter in the soil. Numerous strains of *S. meliloti* have been noted for tolerance of alkaline soil, high salt, and drought conditions (14).

The common model host legume for *S. meliloti* is the barrel medic (*Medicago truncatula*). However, we and others prefer to use the closely related alfalfa (*Medicago sativa*) because it is much more relevant to agriculture (Figure 1.3). *S. meliloti* can also nodulate certain members of the genera *Melilotus* and *Trigonella*. Initial symbiotic signaling begins by the perception of luteolin and chrysoeriol by the bacteria (discussed below). The cognate bacterial signal is a lipochitooligosaccharide with succinyl modifications. *S. meliloti* enters via infection threads, and indeterminate nodules are formed in this partnership (15, 16).

The United States produced over 54 million tons of alfalfa in 2019. This crop is used as hay for ruminants because it has more protein than grass hay and is a better source of calcium and phosphorus (17). In addition, alfalfa is a popular rotation crop because it naturally returns nitrogen to the soil. When the alfalfa senesces or are tilled into the ground, the nitrogen they acquired from

the atmosphere enriches the surrounding soils. In contrast, other crops, including many legumes, require substantial inputs of fertilizers to maximize yields. Only a fraction of synthetic fertilizer goes to crop plants, and a significant portion leaves farmland as runoff, which disrupts the water systems it enters (18).

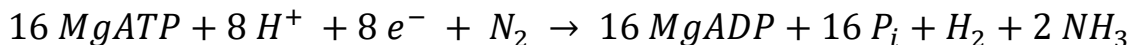
### **The symbiotic interaction between *S. meliloti* and alfalfa**

The gesticulations both plant and bacteria go through to achieve a successful symbiosis are highly involved. As we argue throughout this thesis, the first step in this interaction is the recruitment and localization of the rhizobia. The rhizobia begin dispersed throughout the soil but can make their way to the plant host by following a trail of chemical cues using their chemotaxis systems (discussed later). Only certain sites on the plant root can accept the rhizobia, specifically, young root hairs. It also happens that exudation of nutrient molecules is maximal from these sites. Next, the bacteria must adhere to the surface of the plant by forming biofilms and utilizing proteinaceous adhesive elements (4, 19). At this point, the bacteria are close to sufficiently high concentrations of those compounds that induce genes involved in the nodulation (*nod* genes) process. As stated above, luteolin and chrysoeriol are inducers of *nod* genes in *S. meliloti*, but host-derived stachydrine and trigonelline are also capable of this. Furthermore, the anti-inducers coumestrol and medicarpin, compounds that prevent the induction of *nod* genes in the presence of inducers, are also found in alfalfa seed and root exudates, though their purpose is not well understood (20, 21). The sensors responsible for these compounds are the NodD proteins, and a bacterium may have several, each with a different ligand profile. The role each inducer and anti-inducer compound plays in the chemical interplay between host and microsymbiont is still incompletely understood (8).

Among the primary suites of *nod* genes are the Nod-Factor (NF) biosynthesis genes. NFs are lipochitooligosaccharides that consist of a lipid chain of various unsaturation, 3 to 5 N-acetylglucosamine (chitin) units, and various peripheral modifications to the chitin moieties, including succinylation, sulfonation, and glycosylation (9). The binding of the cognate NF to the LysM-type receptor kinase begins a cascade of signaling events, most salient of which is the curling of the root hair tip. Root hair curling serves to concentrate the chemical signals and provide a positive feedback loop. At the curled tip, parts of the plant cell wall break down as bacteria are engulfed into invaginations. This reconstruction evolves into the infection thread, the tubular

structure that will ultimately deliver the bacteria to the nodule cells. Once inside the infection thread, the bacterium becomes entirely dependent on the plant host for survival (16). Those bacteria that were not engulfed into infection threads can reasonably persist on a continued low-level of root exudates. The infection thread continues to grow towards the cortex of the root where primordial nodule cells have already been differentiating in preparation to receive the rhizobia. Once the nodule is nearly mature, the infection threads, which have ramified, deposit their bacterial cargo via infection droplets (22). These are engulfed by a plant membrane and separate into the beginnings of the symbiosome – the impermanent organelle that is responsible for nitrogen fixation. Inside the symbiosome, the rhizobia have been differentiating from the free-living, reproducing state into the non-reproducing, specialized state called bacteroids (Figure 1.4). This transformation is permanent in determinate nodules and is spurred in part by plant defenses, such as compartment acidification and membrane permeabilization by Nodule-specific Cysteine-Rich peptides (NCR). The plant supplies mineral nutrients, essential amino acids, and malate as a carbon source to the bacteroids (16). Nitrogen fixation can proceed when the bacteroids are fully differentiated. The reaction proceeds as follows in equation 1. Clearly, this is an energy intensive reaction.

Equation 1



The nitrogenase enzyme is unable to differentiate between other substrates such as acetylene, carbon monoxide, and oxygen. Oxygen irreversibly oxidizes and inactivates the enzyme. To combat this, plant cells actively scavenge oxygen from the environment of nodule cells involved in fixation, and the bacteroids rapidly utilize remaining oxygen in respiration (23).

If bacteroids are terminally differentiated, how do the rhizobia pass on and maintain genes for symbiosis and nitrogen fixation? When the infection threads ramify at the cortical cells, some continue to grow just behind the meristem of the growing nodule. The plant will maintain a population of bacteria in these infection threads to seed new nodule cells for as long as the nodule is functional. Eventually, nodules will be replaced or turned over. The surviving bacteria inside infection threads are released during senescence and will have first access to the decaying nodule matter (22).

## **Chemotaxis and motility**

Chemotaxis is the biased movement of an organism towards an attractant or away from a repellent and is made up of a sensory system that communicates with a motility system. Flagellar motility is the system in which a bacterium is propelled by flagella, rotating proteinaceous filaments for which it is named (24, 25). The flagellar filament is made up of flagellin monomers that, combined with the hook and basal body, constitute the entire flagellum. *Escherichia coli* and *S. meliloti* have several flagella distributed evenly about the cell body, an arrangement classified as peritrichous (26).

The chemotaxis system of *E. coli* is the best studied and arguably the simplest, hence its place as the model for bacterial chemotaxis. Sensing of attractants or repellants starts with the diffusion of these molecules into the periplasmic space, where they can be bound by their specific receptor. These receptors are termed Methyl-accepting Chemotaxis Proteins (MCPs) or chemoreceptors. The ligand binding domain of an MCP is typically flanked by two transmembrane domains, which localizes that region to the periplasmic space, while the cytosolic domain of the protein is involved in interacting with the internal signaling complex. Some chemoreceptors lack transmembrane domains, are relegated to the cytoplasm, and sense internal signals (27, 28). Binding of an attractant to the MCP induces a conformational shift in the periplasmic and signaling regions, constituting a movement of information across the inner membrane. This shift is then conveyed via the MCP signaling domain to the CheA protein both directly and through the CheW adaptor protein. Along with its response regulator partner, CheY, CheA makes up the core two-component signaling system of the chemotaxis pathway. CheA is a sensor histidine kinase and is repressed in its autokinase activity by increasing MCP-ligand binding, which prevents it from being able to phosphorylate CheY. With this and the help of constitutive CheZ phosphatase activity, phosphorylated CheY levels are kept low. In *E. coli*, the result is the maintenance of counterclockwise flagellar rotation, which allows for synchronous rotation and bundling of all the flagella, propelling the bacterium forward in what is called a run (28, 29). When attractant binding decreases, or a repellent binding occurs, CheA autophosphorylates on a conserved histidine residue, which then phosphorylates the conserved aspartate residue of CheY at a rate that outpaces the CheZ phosphatase activity. CheY-phosphate interacts with the motor complex through FliM, the switch protein, inducing a switch to clockwise rotation of the flagellum, which then causes the

unbundling of the flagella and random tumbling of the cell to change direction. Bacteria require a mechanism to compare previous states of MCP binding when receiving constant stimuli. This is fulfilled by the CheR/CheB methyl adaptation system. CheR is a methyl transferase that adds methyl groups to several conserved glutamate residues of each MCP at a steady rate, which prevents the inhibition of CheA autokinase activity and increases the sensitivity of the chemoreceptor to signals. CheB is a methylesterase that, when phosphorylated by CheA-phosphate, can remove methyl groups from the MCP. In concert with ligand binding and the transfer of signals through CheA, the methylation states of the MCPs rise and fall following recent signal transduction, or lack thereof. This lag of methylation is how the cell compares its previous condition to its current one as it swims through the medium (29, 30). Together, the methylation system increases the dynamic range of the receptors and allows the bacterium to adapt to constant stimuli. When the cell swims to an area of lower attractant concentration, CheA autophosphorylates and CheY-phosphate levels rise, which induces a tumble so the cell can reorient to a new direction. If the cell senses it is swimming up an attractant gradient, CheA is prevented from autophosphorylating, MCP methylation decreases, and CheY-phosphate levels are kept low, which prolongs the smooth swim toward the source of the attractant. The probability of switching between counterclockwise and clockwise rotation is therefore determined by the presence or absence of ligand binding. This oscillation between runs and tumbles is called a biased, random walk where the net translocation is towards attractants and away from repellents (31, 32).

*S. meliloti* shares most of the core *E. coli* chemotaxis signaling systems, but has multiple distinguishing features (Figure 1.5). Most notably is that *S. meliloti* never rotates its flagella counterclockwise. Tumbles are instead achieved by asynchronous slowing of the flagellar bundle. Fast, synchronous rotation creates runs (33, 34). The *S. meliloti* system lacks a CheZ phosphatase. Instead, signal is propagated and terminated using two separate response regulators, CheY1 and CheY2. CheY2 performs the canonical role of interacting with the motor to affect changes in rotation. CheY1 is also phosphorylated by CheA but does not interact with the motor. CheY2-phosphate can retrophosphorylate CheA, and the phosphate can be passed to CheY1. (35). Another protein absent in the enteric chemotaxis pathway is CheS. This protein binds to and increases the affinity of CheA for CheY1 and thus favors the phosphorylation of CheY1 over CheY2 in what is called a phosphate sink. Spontaneous dephosphorylation of CheY1 and CheY2 removes signal from the system. (36).



*S. meliloti* also possesses a CheD protein in its primary chemotaxis system. The homolog to this protein in *Bacillus subtilis* is a deaminase that converts glutamines into glutamates on MCPs so that they can be methylated in the adaptation system (37). The role CheD plays in the *S. meliloti* chemotaxis system is probably similar, but yet to be determined. Multiple CheWs are encoded in the *S. meliloti* genome. The major chemotaxis operon contains *cheW1*, while *cheW2* is cotranscribed with *mcpW*. CheW3 does not appear to be necessary for chemotaxis (38). A CheT protein is also found in *S. meliloti* as well as related alpha-proteobacteria. It seems to interact with CheR, but the specific function is as yet undiscovered (unpublished results).

The two best studied chemoreceptors are the aspartate and serine receptors, Tar and Tsr in *E. coli*. Both consists of a 4-helix bundle ligand binding domain, transmembrane domains, a HAMP domain (Histidine kinase, Adenylyl cyclase, Methyl-binding proteins, and Phosphatases), and a signaling domain. 4-helix bundle domains are obligate dimers and bind one ligand per dimer (39). The HAMP domain is involved in conversion of signal binding to kinase modulation (40). The signaling domain forms a hairpin turn at the bottom of the receptor and is the coupling site with CheW and CheA (29). Single chemoreceptors form homodimers which then form trimers (Figure 1.6). In *E. coli*, only one dimer pair affects kinase activity, while another dimer is involved in CheW binding. All chemoreceptors in the cell localize to one pole and form a hexagonal chemosensory array with CheA and CheW binding partners. The proximity between the receptors allows signals to be amplified and increases both the sensitivity and signal integration of the sensory system. Cytoplasmic chemoreceptors also form trimers of receptor dimers, but make up arrays that consist of two layers flanked with CheA/CheY plates on either side, and are not necessarily restricted to the cell pole or alongside the array of transmembrane chemoreceptors (27, 30). Tar and other MCPs may use two independent mechanisms to sense very different attractants. Aspartate is bound directly in the canonical binding pocket between the two Tar monomers (42). Maltose sensing is achieved when periplasmic maltose binding protein (MBP) binds maltose and the MBP-maltose complex then interacts with Tar (43). This indirect sensing mechanism broadens the range of compounds a bacterium can sense beyond the specificities of its chemoreceptors (44). *S. meliloti* has eight expressed chemoreceptors (Figure 1.7). Six are transmembrane (McpT, McpU, McpV, McpW, McpX, and McpZ), and two are cytoplasmic (McpY, and IcpA). IcpA (Internal Chemotaxis Protein A) lacks methylation sites found the in the other MCPs. The first step in

chemotaxis is binding of the ligand to the periplasmic region of the MCP (41, 45). These regions are made up of variable domains, such as PAS, Cache, protoglobin, and coiled coil dimers, that bind small molecules. The Cache (Calcium and Chemotaxis) domain is among the most ubiquitous sensor domain in prokaryotes, and also manifests itself as a component in eukaryotic calcium channels. The PAS (Per, Ant, Sim) domain is similar to Cache, but is exclusive to cytoplasmic sensors and lacks an N-terminal alpha-helix found in Cache domains (46). The two cytoplasmic receptors of *S. meliloti*, IcpA and McpY, are found localized to the pole with the other six transmembrane chemoreceptors (41).

Prior to this dissertation, two of *S. meliloti*'s chemoreceptors have been characterized; McpU directly binds most amino acids, and McpX directly binds quaternary ammonium compounds such as glycine betaine, stachydrine, and trigonelline and mediates chemotaxis to these compounds. Both receptors contain dual Cache domains. All the compounds sensed by McpU and McpX were found in host seed exudate (47-49). IcpA is a chemotaxis protein that lacks HAMP domains and has a putative protoglobin domain similar to that found in the oxygen sensor, HemAT, from *B. subtilis* (45, 50). McpY is also cytosolic and has a HAMP domain between its dual PAS domains and signaling domain. Similarity searches indicate it may have an FAD binding domain, which can be used to sense the redox state of the cell (45, 50).

### **Objectives of this Work**

The legumes are unique among most land plants in their symbiotic association. While there are plentiful examples of algae, protists and some land plants that have similar associations, most major crops do not. Engineering a symbiotic system into a crop plant could have a significant impact for farmers that rely on chemical fertilizers. This is not an easy task, so we continue our work to develop a deep understanding of how this symbiosis is initiated in its most nascent state. We begin in chapter 2 by examining the chemotactic potential of flavonoids. These compounds had been deemed chemoattractants by research in the past, but our reexamination reveals that this simply is not the case in our system (51, 52). In ruling out the host-derived flavonoids, we discovered a key clue about the nature of chemoattraction; that nearly all the host-derived attractants are hydrophilic in nature. With that in mind, we turned our focus to the 8 chemoreceptors of *S. meliloti*, as these are the sensors that are directly responsible for seeking our plant hosts. Chapter 3 is a rigorous study of the attractant profile that McpV senses, and an

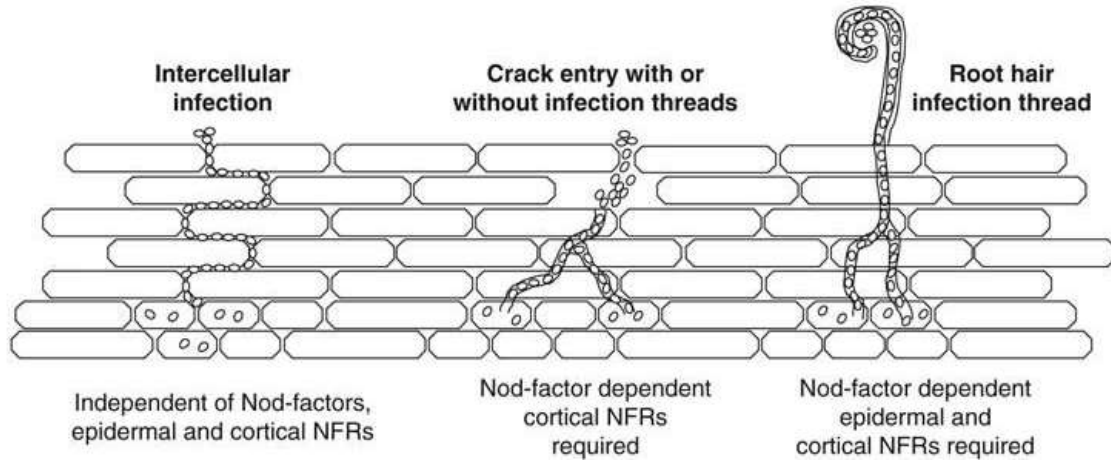
examination of the molecular basis for ligand binding. We determine that short chain monocarboxylates are attractants of *S. meliloti* but find that they are weak compared to the amino acids and may not have a strong influence in host seed sensing. Chapter 4 follows the previous pattern and characterizes McpT as an oxalate sensor and shows its involvement in taxis to di- and tri- carboxylates. This study provides the first evidence of indirect binding in *S. meliloti* chemoreceptors. Finally, chapter 5 is a deep look at the structural basis for the binding of McpX to its ligands, the quaternary ammonium compounds. These works unify into a broadening of our understanding of what signals plants use to recruit bacterial hosts, what chemical cues are of pertinence to bacteria based on their array of sensors, and how these sensors accomplish their purpose.

## REFERENCES

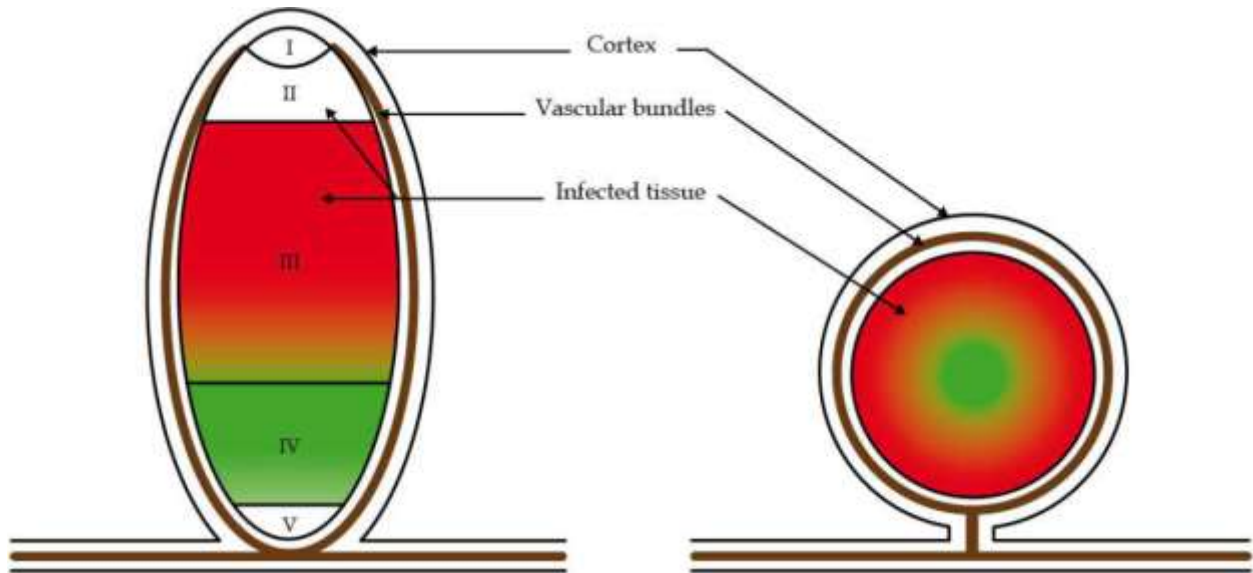
1. Sprent JI. Evolving ideas of legume evolution and diversity: a taxonomic perspective on the occurrence of nodulation. *New Phytol.* 2007;174(1):11-25.
2. Liu S, Ratet P, Magne K. Nodule diversity, evolution, organogenesis and identity. *Regulation of Nitrogen-Fixing Symbioses in Legumes.* 94. London: Elsevier; 2020.
3. Gage DJ. Infection and invasion of roots by symbiotic, nitrogen-fixing rhizobia during nodulation of temperate legumes. *Microbiol Mol Biol Rev.* 2004;68(2):280-300.
4. Wang D, Yang SM, Tang F, Zhu HY. Symbiosis specificity in the legume - rhizobial mutualism. *Cell Microbiol.* 2012;14(3):334-42.
5. Brewin NJ. Development of the legume root nodule. *Annu Rev Cell Biol.* 1991;7:191-226.
6. Gyaneshwar P, Hirsch AM, Moulin L, Chen WM, Elliott GN, Bontemps C, et al. Legume-nodulating betaproteobacteria: diversity, host range, and future prospects. *Mol Plant Microbe Interact.* 2011;24(11):1276-88.
7. Pueppke SG, Broughton WJ. *Rhizobium* sp. strain NGR234 and *R. fredii* USDA257 share exceptionally broad, nested host ranges. *Mol Plant Microbe Interact.* 1999;12(4):293-318.
8. Fisher RF, Long SR. Rhizobium - plant signal exchange. *Nature.* 1992;357(6380):655-60.
9. Oldroyd GED, Murray JD, Poole PS, Downie JA. The rules of engagement in the legume-rhizobial symbiosis. *Annu Rev Genet.* 2011;45:119-44.
10. Abdel-Lateif K, Bogusz D, Hocher V. The role of flavonoids in the establishment of plant roots endosymbioses with arbuscular mycorrhiza fungi, rhizobia and *Frankia* bacteria. *Plant Signal Behav.* 2012;7(6):636-41.
11. Rogel MA, Ormeno-Orrillo E, Martinez Romero E. Symbiovars in rhizobia reflect bacterial adaptation to legumes. *Syst Appl Microbiol.* 2011;34(2):96-104.
12. Young JM. *Sinorhizobium* versus *Ensifer*: may a taxonomy subcommittee of the ICSP contradict the Judicial Commission? *Int J Syst Evol Microbiol.* 2010;60(Pt 7):1711-3.
13. Wibberg D, Blom J, Ruckert C, Winkler A, Albersmeier A, Puhler A, et al. Draft genome sequence of *Sinorhizobium meliloti* RU11/001, a model organism for flagellum structure, motility and chemotaxis. *J Biotechnol.* 2013;168(4):731-3.
14. Mohammad RM, Akhavan-Kharazian M, Campbell WF. Identification of salt- and drought-tolerant *Rhizobium meliloti* L. strains. *Plant Soil.* 1991;134:271-6.
15. Hartwig UA, Maxwell CA, Joseph CM, Phillips DA. Chrysoeriol and luteolin released from alfalfa seeds induce nod genes in *Rhizobium meliloti*. *Plant Physiol.* 1990;92(1):116-22.
16. Udvardi M, Poole PS. Transport and metabolism in legume-rhizobia symbioses. *Annu Rev Plant Biol.* 2013;64:781-805.
17. Higginbotham G, Stull C, Peterson N, Rodiek A, Reed B, Guerrero J. Alfalfa utilization by livestock. University of California Division of Agriculture and Natural Resources; 2008. Contract No.: 8303.
18. Fields S. Global nitrogen: cycling out of control. *Environ Health Perspect.* 2004;112(10):A556-63.
19. Luciana V, Rinaudi WG. An integrated view of biofilm formation in Rhizobia. *FEMS Microbiology Letters.* 2010;304:11.
20. Phillips DA, Joseph CM, Maxwell CA. Trigonelline and Stachydrine Released from Alfalfa Seeds Activate NodD2 Protein in *Rhizobium meliloti*. *Plant Physiol.* 1992;99(4):1526-31.

21. Zuanazzi JAS, Clergeot PH, Quirion JC, Husson HP, Kondorosi A, Ratet P. Production of *Sinorhizobium meliloti* nod gene activator and repressor flavonoids from *Medicago sativa* roots. *Mol Plant Microbe In.* 1998;11(8):784-94.
22. Denison RF, Kiers ET. Life histories of symbiotic rhizobia and mycorrhizal fungi. *Current biology.* 2011;21(18):R775-R85.
23. Hoffman BM, Lukoyanov D, Yang ZY, Dean DR, Seefeldt LC. Mechanism of nitrogen fixation by nitrogenase: the next stage. *Chem Rev.* 2014;114(8):4041-62.
24. Adler J. Chemotaxis in bacteria. *Science.* 1966;153(3737):708-16.
25. Adler J. Chemotaxis in bacteria. *Annu Rev Biochem.* 1975;44:341-56.
26. Götz R, Limmer N, Ober K, Schmitt R. Motility and chemotaxis in 2 strains of rhizobium with complex flagella. *J Gen Microbiol.* 1982;128(Apr):789-98.
27. Bi S, Lai L. Bacterial chemoreceptors and chemoeffectors. *Cell Mol Life Sci.* 2015;72(4):691-708.
28. Salah Ud-Din AIM, Roujeinikova A. Methyl-accepting chemotaxis proteins: a core sensing element in prokaryotes and archaea. *Cell Mol Life Sci.* 2017;74(18):3293-303.
29. Parkinson JS, Hazelbauer GL, Falke JJ. Signaling and sensory adaptation in *Escherichia coli* chemoreceptors: 2015 update. *Trends Microbiol.* 2015;23(5):257-66.
30. Hazelbauer GL, Falke JJ, Parkinson JS. Bacterial chemoreceptors: high-performance signaling in networked arrays. *Trends Biochem Sci.* 2008;33(1):9-19.
31. Vladimirov N, Sourjik V. Chemotaxis: how bacteria use memory. *Biol Chem.* 2009;390(11):1097-104.
32. Micali G, Endres RG. Bacterial chemotaxis: information processing, thermodynamics, and behavior. *Curr Opin Microbiol.* 2016;30:8-15.
33. Sourjik V, Schmitt R. Different roles of CheY1 and CheY2 in the chemotaxis of *Rhizobium meliloti*. *Molecular microbiology.* 1996;22(3):427-36.
34. Attmannspacher U, Scharf B, Schmitt R. Control of speed modulation (chemokinesis) in the unidirectional rotary motor of *Sinorhizobium meliloti*. *Mol Microbiol.* 2005;56(3):708-18.
35. Scharf B, Schmitt R. Sensory transduction to the flagellar motor of *Sinorhizobium meliloti*. *J Mol Microbiol Biotechnol.* 2002;4(3):183-6.
36. Dogra G, Purschke FG, Wagner V, Haslbeck M, Kriehuber T, Hughes JG, et al. *Sinorhizobium meliloti* CheA complexed with CheS exhibits enhanced binding to CheY1, resulting in accelerated CheY1 dephosphorylation. *J Bacteriol.* 2012;194(5):1075-87.
37. Kristich CJ, Ordal GW. *Bacillus subtilis* CheD is a chemoreceptor modification enzyme required for chemotaxis. *J Biol Chem.* 2002;277(28):25356-62.
38. Arapov TD, Saldana RC, Sebastian AL, Ray WK, Helm RF, Scharf BE. Cellular stoichiometry of chemotaxis proteins in *Sinorhizobium meliloti*. *J Bacteriol.* 2020;202(14).
39. Ulrich LE, Zhulin IB. Four-helix bundle: a ubiquitous sensory module in prokaryotic signal transduction. *Bioinformatics.* 2005;21 Suppl 3:iii45-8.
40. Khursigara CM, Wu X, Zhang P, Lefman J, Subramaniam S. Role of HAMP domains in chemotaxis signaling by bacterial chemoreceptors. *Proc Natl Acad Sci U S A.* 2008;105(43):16555-60.
41. Meier VM, Scharf BE. Cellular localization of predicted transmembrane and soluble chemoreceptors in *Sinorhizobium meliloti*. *J Bacteriol.* 2009;191(18):5724-33.
42. Milburn MV, Prive GG, Scott WG, Milligan DL, Yeh J, Jancarik J, et al. Structural changes in a transmembrane receptor - crystal-structures of the ligand domain of aspartate chemotaxis receptor with and without aspartate. *Biochemistry-U.S.* 1992;31(7):2192-.

43. Hazelbauer GL. Maltose chemoreceptor of *Escherichia coli*. J Bacteriol. 1975;122(1):206-14.
44. Glekas GD, Mulhern BJ, Kroc A, Duelfer KA, Lei V, Rao CV, et al. The bacillus subtilis chemoreceptor mcpc senses multiple ligands using two discrete mechanisms. J Biol Chem. 2012;287(47).
45. Meier VM, Muschler P, Scharf BE. Functional analysis of nine putative chemoreceptor proteins in *Sinorhizobium meliloti*. J Bacteriol. 2007;189(5):1816-26.
46. Upadhyay AA, Fleetwood AD, Adebali O, Finn RD, Zhulin IB. Cache domains that are homologous to, but different from PAS domains comprise the largest superfamily of extracellular sensors in prokaryotes. PLoS computational biology. 2016;12(4):e1004862.
47. Webb BA, Hildreth S, Helm RF, Scharf BE. *Sinorhizobium meliloti* chemoreceptor McpU mediates chemotaxis toward host plant exudates through direct proline sensing. Applied and environmental microbiology. 2014;80(11):3404-15.
48. Webb BA, Compton KK, Castañeda Saldaña R, Arapov T, Ray WK, Helm RF, et al. *Sinorhizobium meliloti* chemotaxis to quaternary ammonium compounds is mediated by the chemoreceptor McpX. Molecular microbiology. 2017;103(2):333-46.
49. Webb BA, Compton KK, Del Campo JSM, Taylor D, Sobrado P, Scharf BE. *Sinorhizobium meliloti* chemotaxis to multiple amino acids is mediated by the chemoreceptor McpU. Molecular plant-microbe interactions : MPMI. 2017;30(10):770-7.
50. Schweinitzer T, Josenhans C. Bacterial energy taxis: a global strategy? Arch Microbiol. 2010;192(7):507-20.
51. Caetanoanollés G, Cristestes DK, Bauer WD. Chemotaxis of *Rhizobium meliloti* to the plant flavone luteolin requires functional nodulation genes. J Bacteriol. 1988;170(7):3164-9.
52. Dharmatilake AJ, Bauer WD. Chemotaxis of *Rhizobium meliloti* towards nodulation gene-inducing compounds from alfalfa roots. Applied and environmental microbiology. 1992;58(4):1153-8.



**Figure 1.1** Comparison of host infection methods by which rhizobia enter plant roots. Crack infection is regarded as the ancestral invasion method that is shared with parasites. Intercellular invasion takes advantage of the spaces inbetween cells. Infection threads are the most advanced form of invasion and requires host stimulated degradation of cell walls to facilitate the thread formation. All methods results in the deposition of bacteria into cortical cells. Image used under CC BY-NC-SA 3.0 from *Nat Commun* **1**, 10 (2010). <https://doi.org/10.1038/ncomms1009>  
 ©2010Macmillan Publishers Limited

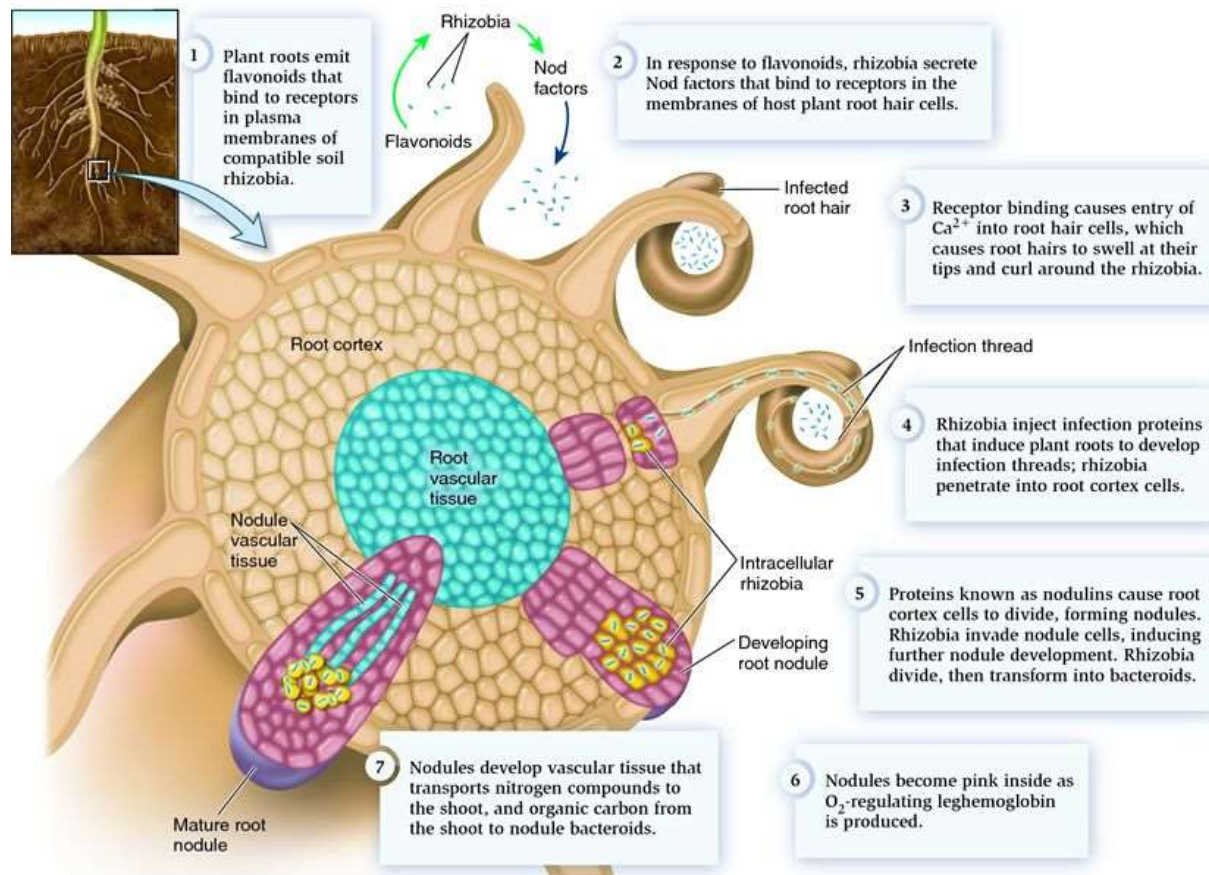


**Figure 1.2** Comparison of indeterminate (left) and determinate (right) nodules. The indeterminate nodule is defined by its persistent meristem and is divided into several zones. I. meristematic zone. II. Infection zone. III. Fixation zone. IV. Senescence zone. V. Saprophytic zone. The nodule originates from and is made up of cortical tissues. Vascular bundles infiltrate the nodule to facilitate metabolite exchange. Image used under creative commons attribution 3.0 license © 2011 Laurence Dupont, Geneviève Alloing, Olivier Pierre, Sarra El Msehli, Julie Hopkins, Didier Hérouart and Pierre Frendo

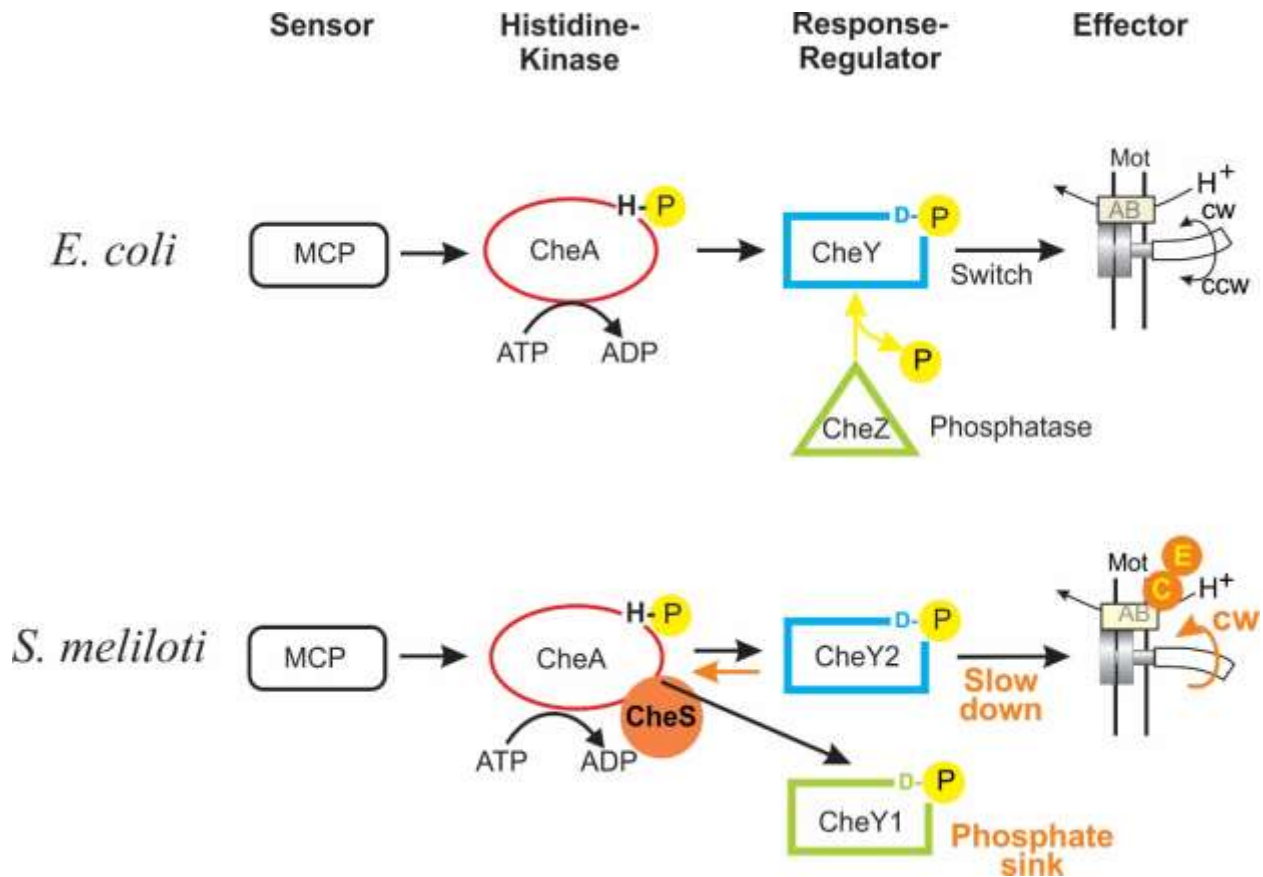




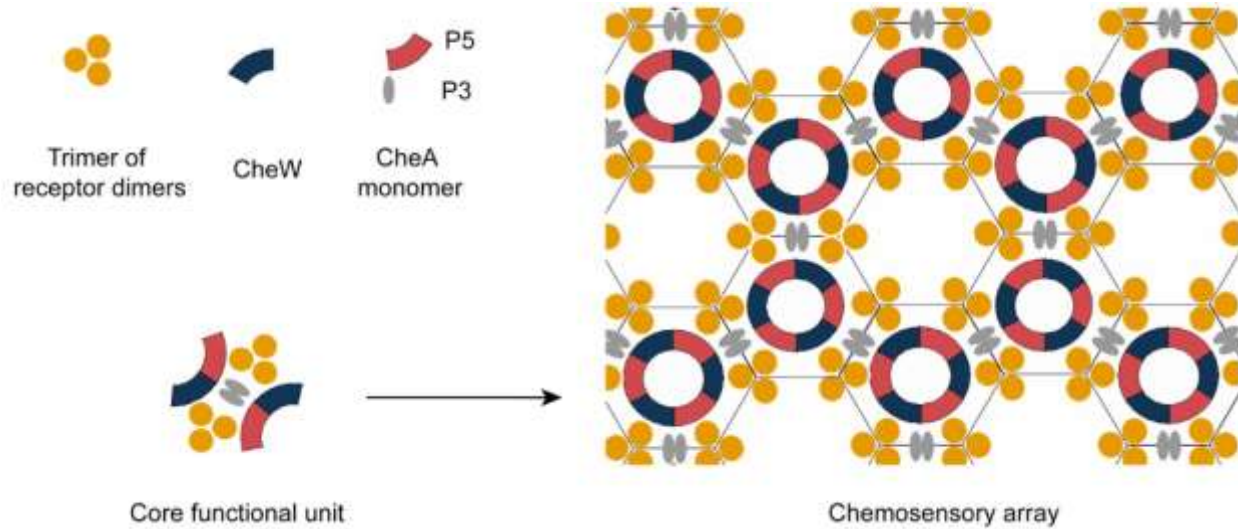
**Figure 1.3** Herbarium specimen of alfalfa (*Medicago sativa* L.). Note the tripinnate leaves and dark, curled seed pods. When mature, the seed pods will split open and scatter the seeds. Used under CC-BY-NC and with permission from the Massey Herbarium, Virginia Tech.



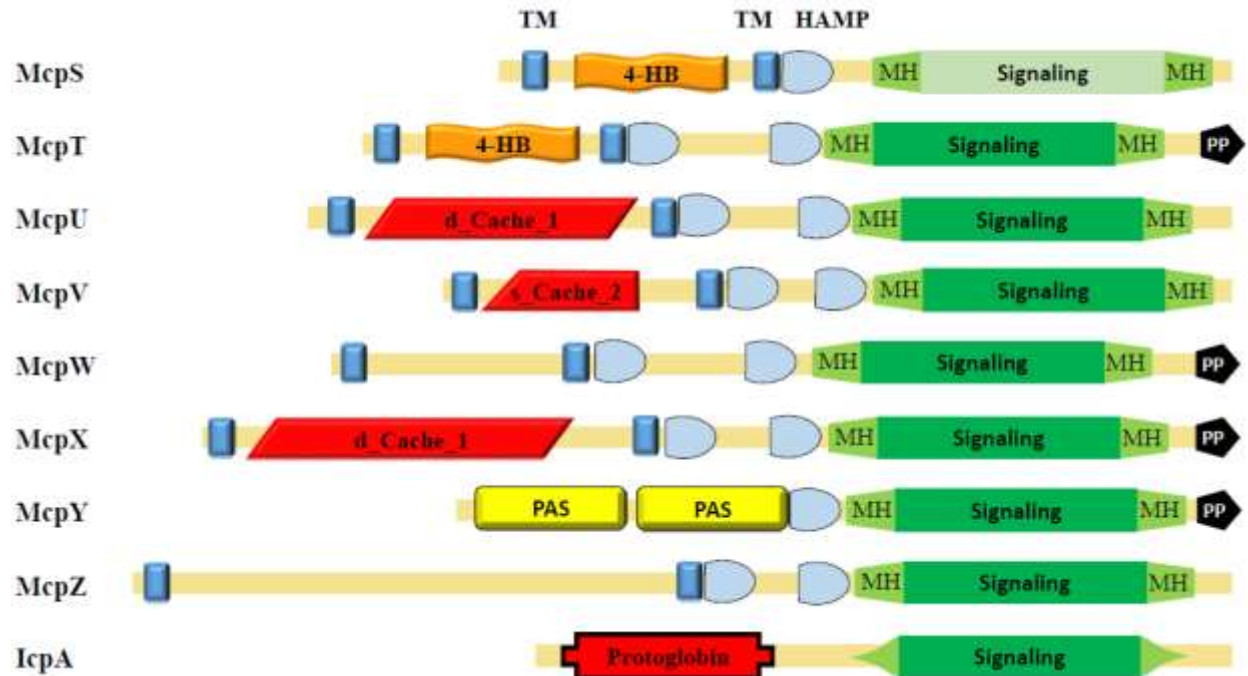
**Figure 1.4** Development of the legume root nodule. Rhizobia and legumes identify their cognate partners by sensing flavonoids and NFs, respectively. Bacteria enter the plant via infection threads while cortical cells divide to accommodate the new organ. Rhizobia enter nodule cells and differentiate into bacteroids where they can begin nitrogen fixation. Used with permission from BiologyForums: [https://biology-forums.com/gallery/33\\_24\\_07\\_11\\_11\\_15\\_13.jpeg](https://biology-forums.com/gallery/33_24_07_11_11_15_13.jpeg)



**Figure 1.5** Comparison of *E. coli* (top) and *S. meliloti* (bottom) chemotaxis systems. In place of the *E. coli* CheZ phosphatase, *S. meliloti* utilizes a second CheY response regulator as a phosphate sink to terminate signal. Used with permission from Birgit Scharf.



**Figure 1.6** Hexagonal chemosensory array consisting of trimers of receptor dimers, CheA, and CheW. This lattice clusters at the pole of the cell and assembles to increase signaling cooperativity and integration. Note that the receptors protrude in and outside of the cell while CheA and CheW are inside the cell. Used with permission from *Cell. Mol. Life Sci.* **72**, 691–708 (2015). <https://doi.org/10.1007/s00018-014-1770-5> © Springer Basel 2014



**Figure 1.7** Domain architecture of *S. meliloti* chemoreceptors. TM: Transmembrane, blue. HAMP: Histidine kinase, Adenylyl cyclase, Methyl-binding proteins, and Phosphatases, light blue. MH: Methylation helix, light green. PP: pentapeptide motif, black. 4-HB: 4-Helix bundle, orange. d\_Cache\_1: Dual Calcium channel and Chemotaxis receptor, red parallelogram. s\_Cache\_1: single Calcium channel and Chemotaxis receptor, red trapezoid. PAS: Per, Arnt, Sim, yellow rectangle. IcpA is classified not classified as an MCP because it lacks methylation helices. McpS is found on an alternant chemotaxis operon and is not expressed when *S. meliloti* is motile.

**Chapter 2 - An updated perspective on *Sinorhizobium meliloti* chemotaxis to alfalfa flavonoids**

K. KARL COMPTON<sup>a</sup>, SHERRY B. HILDRETH<sup>a</sup>, RICHARD F. HELM<sup>b</sup>, BIRGIT E. SCHARF<sup>a</sup>#

<sup>a</sup> Virginia Tech, Department of Biological Sciences, Life Sciences I, Blacksburg, VA, United States of America

<sup>b</sup> Virginia Tech, Department of Biochemistry, Fralin Life Science Institute, Blacksburg, VA, United States of America

Running title: Rhizobial chemotaxis to flavonoids

# For correspondence:

E-mail: bscharf@vt.edu

Keywords: Motility, plant-host exudate, plant-microbe signaling, rhizosphere, symbiosis, solid phase extraction, organic cosolvents

**Attribution: KKC performed chemotaxis experiments and harvested exudate samples. SBH performed analysis of exudate samples. Manuscript was written by KKC, BES, and RFH. Figures 5.1 – 5.6 were made by KKC. Figures 5.7 – 5.10 were made by RFH.**

## ABSTRACT

The symbiotic interaction between leguminous plants and their cognate rhizobia allows for the fixation of gaseous dinitrogen into bioavailable ammonia. The perception of host-derived flavonoids is a key initial step for the signaling events that must occur preceding the formation of the nitrogen-fixing organ. Past work investigating chemotaxis - the directed movement of bacteria through chemical gradients - of *Bradyrhizobium japonicum*, *Rhizobium leguminosarum*, and *Rhizobium meliloti* discovered chemotaxis to various organic compounds, but focused on chemotaxis to flavonoids because of their relevance to the symbiosis biochemistry. The current work sought to replicate and further examine *Sinorhizobium (Ensifer) meliloti* chemotaxis to the flavonoids previously thought to act as the principal attractant molecules prior to the initial signaling stage. Exudate from germinating alfalfa seedlings was analyzed for composition and quantities of different flavonoid compounds using mass spectrometry. The abundance of four prevalent flavonoids in germinating alfalfa seed exudates was at a ratio of 200:5:5:1 for hyperoside, luteolin, luteolin-7-glucoside, and chrysoeriol. Using quantitative chemotaxis capillary assays, we did not detect chemotaxis of motile *S. meliloti* cells to these and two other flavonoids identified in seed exudates. In support of these findings, the flavonoid fraction of seed exudates was found to be an insignificant attractant relative to the more hydrophilic fraction. Additionally, we observed that cosolvents commonly used to dissolve flavonoids confound the results. We propose that the role flavonoids play in *S. meliloti* chemotaxis is insignificant relative to other components released by alfalfa seeds.

## INTRODUCTION

Plants of the Fabaceae family share a unique association with specific soil-dwelling bacteria that allows the plants access to the otherwise unavailable nitrogen in the atmosphere. This association with their cognate bacterial symbionts, referred to as rhizobia, has precipitated a great diversification of species in this family, yielding some of the most important crops in agriculture (Brewin, 1991; van Rhijn and Vanderleyden, 1995; Sprent, 2007).

Over the last several decades, the molecular basis for the development of this interaction has been rigorously expounded. The first step initiating this symbiosis is the release of flavonoids by the plant. Flavonoids are phenylpropanoid derivatives and act as signaling molecules, antimicrobials, and growth promoters (Waage and Hedin, 1985; Peters et al., 1986; Peters and Long, 1988; Hartwig et al., 1990a; Hartwig et al., 1991; Wachter et al., 1999; Sato et al., 2000). Rhizobia receive these flavonoid signals using the sensor protein NodD, which, subsequent to proper flavonoid binding, will induce the expression of “*nod*-box genes” (Peters et al., 1986; Hartwig et al., 1990b; a; Peck et al., 2006). Each rhizobial species may have one or several NodD copies, each with a unique specificity for certain flavonoids (Fisher and Long, 1992; Phillips et al., 1992). While a single flavonoid is sufficient to induce *nod* gene expression, rhizobia are exposed to a cocktail of host-derived flavonoids (Maxwell et al., 1989; Hartwig et al., 1990a; Maxwell and Phillips, 1990). Within a species, the makeup of this cocktail varies dramatically depending on the growth stage, location, and status of the plant (Maxwell et al., 1989; Maxwell and Phillips, 1990).

Flavonoids are not the only host-derived compounds that can affect *nod* gene expression. For example, trigonelline and stachydrine from *Medicago spp.* can induce *nod* gene expression, while the isoflavonoid-derivative medicarpin and the phytoestrogen coumestrol are both antagonists of *nod* gene expression (Phillips et al., 1992; Zuanazzi et al., 1998). The total involvement of these compounds in the symbiotic interaction is incompletely understood. Among the suite of genes induced by NodD are those involved in the synthesis of Nod factors, which are lipochitooligosaccharides that reciprocate the symbiotic signal to the plant. Reception of Nod factors occurs at the root hair. Among the most salient changes is the curling of the root hair around the population of rhizobia and the formation of the infection thread (IT), an invagination of the plant cell that will traverse into the cortical cells of the root. The rhizobia occupy the IT as it makes its way to what will become nodule primordia. The IT will ramify and act to hold a population of rhizobial cells that will terminally differentiate into nitrogen-fixing bacteroids once seeded into



nodule cells (Brewin, 1991; van Rhijn and Vanderleyden, 1995; Wang et al., 2012; Haag et al., 2013; Udvardi and Poole, 2013).

Prior to all the above, the rhizobia must localize to the tips of developing root hairs. Because the nodule is a highly sought-after niche, there is strong competition among rhizobia that can successfully colonize the host (Pinochet et al., 1993; Hirsch and Spokes, 1994; Da and Deng, 2003). While the flagella-driven motility and chemotaxis system has been shown to be unnecessary for nodule formation and nitrogen fixation, motility and chemotaxis provide a competitive advantage in all investigated rhizobia-host symbioses. In the rhizosphere, locomotion and navigation create the ability to out compete neighbors for nodule occupancy (Ames et al., 1980; Napoli and Albersheim, 1980; Ames and Bergman, 1981; Miller et al., 2007; Lacal et al., 2010; Scharf et al., 2016). Motility and chemotaxis are critical for bacteria that occupy multiple niches or need to seek out spatial niches (Lacal et al., 2010). Thus, chemotaxis systems are ubiquitous in numerous clades of rhizobia.

Chemotaxis is the systematic movement of an organism through a chemical gradient, and in bacteria operates on the principle of a biased random walk. Bacteria will rotate their flagella to either swim smoothly in a largely straight direction or tumble in place to reorient. The frequency at which cells switch between both modes is dependent on their movement up or down a chemical gradient. If cells are swimming down an attractant gradient (*i.e.*, away from the source of the attractant), they will initiate tumbles more frequently to reorient themselves in the proper direction. If cells are swimming up an attractant gradient (*i.e.*, towards the source of the attractant), the tumbling behavior is suppressed and smooth swimming is protracted. The net effect of this precept is a gradual translocation to the source of an attractant molecule (Berg, 2003; Parkinson et al., 2015).

The first examination of chemotaxis in *Sinorhizobium (Ensifer) meliloti* (then called *Rhizobium meliloti*) was in a 1976 study by Currier and Strobel (Currier and Strobel, 1976). They tested six different strains of rhizobia for chemotaxis to host legume, non-host legume, and non-legume root exudates. Later, in the 1980s, most publications on the chemotaxis of *S. meliloti* noted taxis to both sugars and amino acids, with the latter appearing to be preferred (Burg et al., 1982; Götz et al., 1982; Malek, 1989). In addition, select studies examined the possibility of chemotaxis to *nod* gene-inducing flavonoids (Caetanoanollés et al., 1988; Dharmatilake and Bauer, 1992). However, these studies lacked context in the form of other classes of test compounds and presented what could be

considered comparatively low levels of chemotaxis. Since then, great knowledge was gained about rhizobial chemotaxis, components of the signaling system, and molecular signaling mechanisms (Pleier and Schmitt, 1991; Sourjik and Schmitt, 1996; Sourjik et al., 2000; Rotter et al., 2006). However, follow-up experiments on flavonoid chemotaxis are lacking and the mechanism of flavonoid sensing in chemotaxis remains undefined.

At outset, the aim of this work was to identify the chemotactic sensor for flavonoids in *S. meliloti*. The most common and best studied mechanism of chemotactic sensing is the direct binding of an attractant molecule to the periplasmic region of a Methyl-accepting Chemotaxis Protein (MCP). As of now, our lab has identified the ligand classes for three of the eight *S. meliloti* MCPs known to be involved in chemotaxis (Meier et al., 2007; Meier and Scharf, 2009; Webb et al., 2014; Webb et al., 2017a; Webb et al., 2017b; Compton et al., 2018). The original goal of this study was particularly salient, especially since, to our knowledge, this would have been the first documented example of a chemotactic sensor for flavonoids in bacteria. We first identified the main flavonoids present in germinating seed exudates of the economically relevant host, alfalfa (*Medicago sativa*). Next, we tested commercially available standards of these compounds in chemotaxis assays. As we tested multiple compounds and experimental conditions, it became apparent that *S. meliloti* was not attracted to any flavonoids identified in alfalfa seed exudates, and therefore, could not confirm previous studies. Based upon these results, we reexamined the relevance of flavonoid chemotaxis in the recruitment of symbiotic rhizobia to the roots of their legume hosts.

## MATERIALS AND METHODS

**Chemotaxis assays.** Assays for chemotaxis were done in a manner derived from the capillary-based method of Adler (Adler, 1973). Motile *S. meliloti* RU11/001 cells were grown by inoculating 10 ml of Rhizobium Basal medium (RB: 0.1 mM NaCl, 10  $\mu$ M Na<sub>2</sub>MoO<sub>4</sub>, 6.1 mM K<sub>2</sub>HPO<sub>4</sub>, 3.9 mM KH<sub>2</sub>PO<sub>4</sub>, 1 mM (NH<sub>4</sub>)<sub>2</sub>SO<sub>4</sub>, 1  $\mu$ M FeSO<sub>4</sub>, 1 mM MgSO<sub>4</sub>, 0.1 mM CaCl<sub>2</sub>, 20  $\mu$ g/l D-biotin, and 10  $\mu$ g/l thiamine) overlain on a Bromfield medium agar plate (0.4 g/l tryptone, 0.1 g/l yeast extract, 0.45 mM CaCl<sub>2</sub>, and 15 g/l agar) (Götz et al., 1982; Sourjik and Schmitt, 1996) overnight at 30 °C. Cells were harvested at an OD<sub>600</sub> of 0.15 to 0.18, centrifuged at 3,000 x g for 5 minutes to remove spent culture, and suspended to an OD<sub>600</sub> of 0.15. After checking microscopically that greater than 50 % of the population were motile, 350  $\mu$ l of culture was dispensed into a flat glass chemotaxis well. One-microliter microcaps (Drummond Scientific) were sealed at one end and filled with solution using a vacuum or centrifugation. Chrysoeriol, hyperoside, luteolin-7-O-

glucoside, and quercetin-3-O-(6''-acetylglucoside) were acquired from Extrasynthese (Genay, France), quercetin and luteolin from Cayman Chemical (Ann Arbor, MI), and pratensein from Chromadex (Irvine, CA). Flavonoids were dissolved in 100 % methanol or dimethyl formamide (DMF) and appropriately diluted in RB. To prepare hyperoside in the absence of a cosolvent, an aliquot was suspended in RB by vortexing and brief incubation at 42 °C. After centrifugation, the hyperoside concentration was determined using the molar extinction coefficients  $\epsilon_{259} = 20400 \text{ M}^{-1}\text{cm}^{-1}$  and  $\epsilon_{364} = 24500 \text{ M}^{-1}\text{cm}^{-1}$  (Windholz and Merck & Co., 1983). Capillaries were placed in the chemotaxis wells and left to incubate at room temperature for 2 h. Assays were performed for each concentration in technical triplicate for each of three biological replicates, excepting the methanol dose response experiment, which was done in technical duplicate for each of four biological replicates. The capillaries were broken at the sealed end and their contents were dispensed into RB and appropriately diluted. Dilutions were plated onto tryptone, yeast, calcium chloride plates (TYC: 5 g/l tryptone, 3 g/l yeast extract, 5.9 mM CaCl<sub>2</sub>, and 15 g/l agar) with 0.6 mg/ml streptomycin sulfate. Cells per capillary were calculated by subtracting the number of bacteria that accumulated in a capillary with only RB from each test capillary. Alternatively, data are also displayed as chemotaxis ratios, which is the quotient of the cells in the test capillary divided by the cells in the reference capillary. This value was included for easier comparison to previous reports using this method. A capillary containing 10 mM proline or 1 mM sodium acetate was used as a positive control alongside the experiments.

**Flavonoid quantification from germinating seeds.** Seed exudates were harvested from Guardsman II variety alfalfa (*Medicago sativa* L.). For each replicate, 0.1 g seeds were surface sterilized by rinsing four times with sterile water, soaking in 8 ml of 3 % H<sub>2</sub>O<sub>2</sub> for 12 min, and rinsing four times with sterile water. Seeds were left to germinate in 3 ml of sterile water for 24 h at 30 °C. At the time of harvesting, an aliquot of seed exudate was examined for contamination microscopically and plated onto TYC. Samples that did not show contamination in the sample or on the plate the next day were flash frozen in liquid nitrogen and stored at -80 °C.

For solid phase extraction (SPE), 2.5 ml of ten separate seed exudate samples were applied to 1cc Oasis PriME HLB SPE cartridges (Waters, Milford, MA). Each cartridge was washed twice with 1 ml of water and eluted with two 1-ml aliquots of methanol. Flow-through and wash fractions were combined to create the hydrophilic fraction, while the methanol elutions were combined to create the hydrophobic fraction. Both fractions were concentrated to dryness and stored at -20 °C.

For capillary assay experiments and mass spectrometry analysis, both fractions were suspended in water to 5-times their original concentration. When used for capillary assays, fractions and raw exudates were mixed with water and 5-fold concentrated RB to achieve a final experimental concentration of 0.8-fold exudate and 1-fold RB. Since the raw, unfractionated exudates could not be concentrated, 0.8-fold was the highest concentration of raw exudates that could be utilized.

**Mass spectrometry of seed exudate for flavonoid profiling.** Seed exudates were prepared for analysis by dilution in methanol with 0.1 % formic acid (1:1 v/v), sonication in a water bath for 10 min and centrifugation at 13,000 x g for 10 minutes. Sample analysis was performed on a Synapt G2-S high resolution Q-TOF mass spectrometer (Waters Corp., Milford, MA) interfaced with an Acquity I-class UPLC (Waters Corp., Milford, MA). Mobile phases were 0.1 % formic acid (A) and 0.1 % formic acid in acetonitrile (B). The flow rate was 0.2 ml/min and the 20-min elution gradient was: initial 1 % B, 0.5 min hold 1 % B, gradient to 40 % B 12 min; gradient to 90 % B 17.5 min, 18 min hold 90 % B and 19 min return to initial conditions. Two  $\mu$ l of sample was injected onto a Waters BEH C<sup>18</sup> 1.7  $\mu$ m, 50 x 2.1 mm column (Waters Corp., Milford, MA) held at 35 °C. The mass spectrometer was operated in negative mode under high resolution and MS<sup>E</sup> settings with a mass scan range of 50-1800. Instrument parameters were capillary voltage 1.5 kV, source temperature 125 °C, Sampling cone 30 V, Source offset 80, desolvation temperature 350 °C, desolvation gas 500 L/hr, cone gas 50 L/hr and nebulizer gas 6 bar. The cycle time was 0.2 sec and collision energy was set at 4 eV for low energy scans and ramped from 20-40 in the high energy scans. Leucine enkephlan (Waters Corp., Milford, MA) was continuously infused into the source at 5  $\mu$ l/min and analyzed at 20 sec intervals for real-time mass correction.

Data visualization and analysis was performed with MassLynx v 4.2 (Waters Corp., Milford, MA). Peaks corresponding to potential flavonoids were identified in the high energy scan data by searching spectra for aglycone masses related to known flavonoids. Low energy scan data was then used to determine the precursor species and tentative identifications assigned based upon literature and database searches. Authentic standards were purchased and analyzed with the conditions previously described. Assignment of flavonoid identity was based upon standards and seed exudate providing the same mass, retention time, and high energy mass fragments. Analysis of seed exudate fractions following SPE was performed in the same manner as described above for flavonoid profiling.

**Mass spectrometry of seed exudates for flavonoid quantification.** Quantification of flavonoids was performed on a Shimadzu 8060 triple quadrupole mass spectrometer (Shimadzu Corp., Kyoto, Japan) interfaced with a Shimadzu Nextera UPLC (Shimadzu Corp., Kyoto, Japan). The flow rate was 0.4 ml/min and the gradient composition was as follows: initial 40 % B, gradient from 0.5 - 4 min 90 % B, 5 min return to initial composition. Five  $\mu$ l were injected onto a Waters BEH C<sub>18</sub> 1.7  $\mu$ m, 2.1 x 50 mm (Waters Corp., Milford, MA) held at 40 °C. The mass spectrometer was operated in positive ionization and MRMs were developed based upon compound specific transitions (Table 1). Standards were analyzed at concentrations from 1-1000 ng/ml to generate calibration curves. Data were analyzed with Lab Solutions software v 5.93 (Shimadzu Corp., Kyoto, Japan).

## RESULTS

**Identification and quantification of flavonoids in alfalfa seed exudates.** To identify abundant and symbiotically relevant flavonoid species in germinating alfalfa (*Medicago sativa*) seed exudates, we utilized a purification scheme (C<sub>18</sub> solid phase extraction) reported previously for alfalfa seed exudates (Hartwig et al., 1990a). Five flavonoids were identified in this prior work, all related to luteolin, namely chrysoeriol (3'-methoxyluteolin), luteolin, luteolin-7-O-glucoside (cynaroside, L-7-G), 5-methoxyluteolin, and 3'-5-dimethoxyluteolin (Fig 2.7). Trace amounts of apigenin and 4'-7-dihydroxyflavone were also reported. This purification scheme was coupled to a metabolomics profiling platform (ultra-performance liquid chromatography-quadrupole time of flight mass spectrometry, UPLC-QTOF MS) to assess the flavonoid profile in detail.

Chrysoeriol, luteolin, and L-7-G were confirmed to be present in the seed exudates based upon comparison to authentic standards. The other methoxylated luteolins (3',5-dimethoxyluteolin and 5-methoxyluteolin), were tentatively identified based upon the parent mass and reported fragmentation patterns (7) but were not confirmed with authentic standards. The most abundant peak observed in the exudate LC-MS chromatograms was hyperoside (quercetin 3-O- $\beta$ -D-galactopyranoside), which exhibited an identical retention time and fragmentation pattern to an authentic standard.

We next focused on several unknowns that contained fragmentation patterns indicative of flavonoids. Quercetin was found to be present in trace amounts, as was an apparent acetylated quercetin glycoside with a [M-H]<sup>-</sup> mass of 505.0977 *m/z*. Quercetin-3-O-(6-acetylglucoside) was

a possible candidate, but while the fragmentation patterns were similar, the retention times did not match, leaving the compound as a putative acetylated hexosyl flavonoid. While we confirmed the presence of trace levels of apigenin, we did not detect 4'-7-dihydroxyflavone.

The five flavonoid species we have confirmed in seed exudates, namely chrysoeriol, hyperoside, luteolin, L-7-G, and quercetin do not represent the entire flavonoid profile of alfalfa seed exudates. However, these five compounds are representative of the total flavonoid pool, and can be classified by their functional groups into flavones (luteolin, chrysoeriol) and flavonols (quercetin) and their glycosylated variants (L-7-G and hyperoside) (Fig. 2.7). Inasmuch, we next quantified the five identified flavonoids in seed exudates on a per seed basis. Hyperoside was the most abundant flavonoid, followed by luteolin, luteolin-7-glucoside and chrysoeriol. Quercetin was at or below the limit of quantification (Table 1). The ratio of hyperoside, luteolin, L-7-G, and chrysoeriol in alfalfa seed exudates on a per seed basis, was determined to be 200:5:5:1.

***S. meliloti* did not exhibit chemotaxis to alfalfa-derived flavonoids and is unaffected by cosolvents.** The capillary assay is the gold standard for the quantification of bacterial chemotaxis responses. The five flavonoids detected in seed exudates along with an additional isoflavone (pratensein) were tested in the capillary assay using motile cells of the *S. meliloti* wild-type strain RU11/001. We chose growth conditions that had been determined previously to result in optimal motility and expression of the chemotaxis machinery in *S. meliloti* (Rotter et al., 2006; Meier et al., 2007). Methanol was utilized at a final concentration of 4 % (v/v), because the hydrophobic nature of the flavonoids required an organic cosolvent. Therefore, a control with 4 % methanol in RB was also tested. Methanol at this concentration did not affect the motility of *S. meliloti* cells for the duration of the experiments, as evaluated by microscopic observation.

First, we assessed the major *nod*-gene inducing flavonoid luteolin at six concentrations between  $10^{-4}$  and  $10^{-10}$  M based on concentrations used in previous reports of chemotaxis to this compound (Caetanoanollés et al., 1988; Dharmatilake and Bauer, 1992). The number of cells in capillaries containing luteolin did not differ significantly from the control capillary (Fig 2.1). The same lack of chemotaxis was observed for four additional flavonoids at three different concentrations between  $10^{-4}$  and  $10^{-8}$  M (Fig 2.2 A-D). Chrysoeriol could not be assayed using the cosolvent methanol because of its poor solubility in this solvent. Therefore, we decided to use dimethylformamide (DMF) at a final concentration of 2 % (v/v) as an alternative. In addition, we repeated the capillary assays with quercetin using DMF as cosolvent. Similar to the previous

results, a small, but insignificant positive chemotaxis response was observed to the various concentrations of chrysoeriol and quercetin, however, they were indistinguishable from the cosolvent control (Fig 2.3 A, B). In conclusion, we did not observe chemotaxis to any of these flavonoids at any of the tested concentrations.

To further assess the effect of the cosolvent methanol, we examined the chemotaxis of *S. meliloti* to one of its strongest and best characterized attractants, proline, in the absence and presence of 4 % methanol (Meier et al., 2007; Webb et al., 2014). The capillary assays clearly demonstrated that addition of methanol at a final concentration of 4 % did not significantly change the migration of bacteria into the capillary filled with 10 mM proline (Fig 2.4 A). We next considered the possibility that methanol itself served as a weak attractant and tested this by performing a dose-response curve of chemotaxis to methanol between 1 M (approximately 4 % by volume at room temperature) and  $10^{-9}$  M. Methanol did not elicit a chemotaxis response above background at any of the concentrations tested (Fig 2.4 B).

The only flavonoid detected in alfalfa seed exudates that is soluble in water without the aid of an organic cosolvent is hyperoside, a galactoside of quercetin. We took advantage of this property to obviate the use of a cosolvent in the experiments. No chemotaxis to hyperoside was observed in the absence of a cosolvent (Fig 2.5). To assess the possibility that the chemotactic response to a flavonoid is inducible, we grew *S. meliloti* cells in the presence of 5 % alfalfa seed exudate and examined chemotaxis to hyperoside without methanol. The final concentration of hyperoside, the most abundant flavonoid in alfalfa seed exudates, in the growth medium was approximately 4.7  $\mu$ M. Despite this amendment, no chemotaxis to hyperoside was detected (Fig 2.5).

***S. meliloti* is attracted to the seed exudate fractions that are depleted in flavonoids.** We wanted to consider taxis to host derived flavonoids in a more direct context pertaining to seed exudates. To achieve this, we harvested seed exudates and fractionated using C18 solid phase extraction (SPE) cartridges essentially as described previously (Hartwig et al., 1990a). Flow-through and wash fractions with water were collected and pooled into a hydrophilic, “non-flavonoid” fraction. Adsorbed compounds were eluted with 100 % methanol, referred to as the hydrophobic fraction. Confirmation by mass spectrometry analysis showed flavonoids were enriched in the hydrophobic fraction compared with the hydrophilic fraction, which is likely comprised of amino acids, quaternary ammonium compounds (QACs), and carboxylates (Fig. 5.8) (Webb et al., 2017a; Webb et al., 2017b; Compton et al., 2018).

Next, we compared chemotaxis to these two fractions and to the raw, unfractionated exudates. The unfractionated exudates, the hydrophilic fraction, and the combined hydrophilic and hydrophobic fractions all drew approximately 770,000 cells per capillary. The hydrophobic fraction only attracted 26,000 cells per capillary (Fig 5.6), clearly demonstrating that the dominant chemoattractants are hydrophilic in nature and not flavonoids or other hydrophobic compounds.

## DISCUSSION

**A retrospective on the history of rhizobial chemosensing.** Flavonoids are the key chemical signal in the initiation of the symbiosis between legumes and their rhizobial symbionts. This process was conflated with chemotaxis when flavonoids were postulated to also interact with the chemosensory pathways. The first investigation into the roles these compounds play as chemoattractants in *S. meliloti* was reported in 1988, where it was claimed that *S. meliloti* was attracted to luteolin maximally at a concentration of  $10^{-8}$  M, and that this response disappeared when a large portion of the *nif-nod* region was deleted or individual *nod* genes were interrupted by transposon insertion (Caetanoanollés et al., 1988). Reported values were chemotaxis coefficients of approximately 2; an extremely weak chemotaxis response (Götz et al., 1982). In addition, none of the chemotaxis response curves showed the standard error or confidence intervals, making it difficult to compare these responses to the background. Furthermore, the authors did not attempt to recover the wild-type response by complementation (Caetanoanollés et al., 1988). A follow-up study examined the chemotaxis to luteolin, two other host-derived flavonoids, and a chalcone. Unfortunately, and as stated by the authors, the presented data were inconsistent and varied on a day-to-day basis. Nevertheless, *S. meliloti* was concluded to exhibit a chemotaxis to these compounds, though how robust this response was might not have been fully appreciated at the time. (Dharmatilake and Bauer, 1992).

The chemotaxis response to flavonoids was also investigated in other rhizobial species. Aguilar *et al.* examined the chemotaxis of *Rhizobium leguminosarum* bv. *phaseoli* to a number of sugars and phenolics compounds in addition to the flavonoids luteolin, apigenin, and naringenin (Aguilar et al., 1988). These experiments demonstrated chemotaxis to luteolin and apigenin, and the response to these compounds was comparable in magnitude to the response to xylose and several phenolic compounds. Importantly, it was clear from these experiments that the responses were significantly above background (Aguilar et al., 1988). Using *Bradyrhizobium japonicum* as a model organism, Barbour *et al.* tested the contribution of numerous compounds identified in soybean exudates,



including flavonoids to chemotaxis (Barbour et al., 1991). This study revealed that *B. japonicum* is most attracted to carbon sources such as succinate, glutamate, and malonate, but not to the flavonoids luteolin, daidzein, daidzin, and genistein. The effect of transposon insertions in several *nod* genes on chemotaxis to genistein and seed exudates was tested, but no change was observed (Barbour et al., 1991). The lack of chemotaxis to flavonoids in *B. japonicum* corroborated the findings of Kape *et al* (Kape et al., 1991).

While the above studies appear contradictory in nature, only the work of Barbour *et al.* incorporated multiple attractants and utilized relatively stringent statistics. Based on experiments using fractionated seed exudates, the authors aptly concluded that “the primary chemotactic components and the primary [*nod* gene] inducing components are chemically separate” (Barbour et al., 1991).

As it stands, the chemotaxis of rhizobia to flavonoids is widely accepted, stated in numerous reviews and textbooks (Subramanian et al., 2007; Maj et al., 2010; Oldroyd et al., 2011; Abdel-Lateif et al., 2012; Hassan and Mathesius, 2012; White et al., 2012; Liu and Murray, 2016). This is an appealing conclusion given that one or a few highly specific molecules could be responsible for both the recruitment of rhizobial symbionts and induction of their symbiotic pathways. In hindsight, if these studies compared the chemotactic potency of flavonoids to other attractants such as amino acids, flavonoids would not be regarded as significant contributors to the recruitment of rhizobia to host plants.

**A second look at chemotaxis to flavonoids.** Our current study provides an in-depth analysis of the importance of host-derived and symbiotically relevant flavonoids in the attraction of *S. meliloti* to its plant host (Table 1). We performed concentration-dependent chemotaxis assays with *S. meliloti* using four flavonoid aglycones (chrysoeriol, luteolin, pratensein, and quercetin) and two single hexosyl glycones (hyperoside and luteolin-7-glucoside) and did not detect chemotaxis to any of these compounds. Of principal importance is the comparison of our results with previous data on luteolin chemotaxis, as it is the only flavonoid tested by all previous reports (Aguilar et al., 1988; Caetanoanollés et al., 1988; Barbour et al., 1991; Dharmatilake and Bauer, 1992). While we observed similar chemotaxis ratios to previous reports (about 2-3), the negative control unequivocally indicates that this is not due to the luteolin (Fig 2.1) (Caetanoanollés et al., 1988; Dharmatilake and Bauer, 1992). The values of chemotaxis we measured to flavonoids were above the background (reference capillaries containing only buffer), but not distinguishable from the

methanol or DMF cosolvent control (Figs. 2.1-2.3). Since chemotaxis to 10 mM proline was not inhibited by the addition of methanol, the cosolvent does not appear to interfere with the bacterium's ability to sense and swim to attractants (Fig 2.4). Methanol at a concentration of 4 % attracts a certain number of cells but is not an attractant at any concentration below that (Fig 2.4B). As to the cause of the chemoattraction to methanol and DMF, we propose that at approximate concentrations of 1.2 and 0.27 M, respectively, these solvents perturb the chemotaxis signaling system, potentially via membrane disruption (Vaknin and Berg, 2006). We next sought to address the possibility of synergism between flavonoids. To obtain a cocktail that best mirrors what the rhizobia would encounter in the presence of a host, we performed a fractionation of seed exudates, obtaining a hydrophilic fraction and a hydrophobic, flavonoid-containing fraction. Chemotaxis assays to this flavonoid-enriched sample only showed a modest accumulation of cells, similar to that obtained with individual flavonoids. These data also revealed that the hydrophilic fraction is responsible for 100 % of the chemotactic potential of raw seed exudates (Fig 2.6). This information serves as evidence that the best chemoattractants are water-soluble, or at least poorly retained on a reversed phase SPE unit with water as the eluent. Assays were performed with *S. meliloti* cells grown under optimal motility and chemotaxis conditions and were the same as those used to characterize the mechanisms of chemotaxis to amino acids, QACs, and small monocarboxylates (Webb et al., 2014; Webb et al., 2017a; Compton et al., 2018). Examinations of the regulation of motility and chemotaxis in *S. meliloti* allowed us to identify optimal conditions for this behavior to be examined, such as culturing methods, media, growth phase, and cell density (Sourjik et al., 2000; Rotter et al., 2006; Meier et al., 2007). These conditions were used for all subsequent characterizations of *S. meliloti* chemotaxis, including this study. It should be mentioned here that it is practically impossible to prove a negative. Directly put, while we do not confirm any evidence of chemoattraction to flavonoids, this does not disprove the existence of chemotaxis to flavonoids overall, in other bacterial species, under other conditions, or using mechanisms other than flagellar motility. That in mind, we feel that the following lines of evidence make this phenomenon less feasible than it initially appears.

**Evidence against the possibility of flavonoid chemotaxis in *S. meliloti*.** The fact that flavonoid aglycones are at best sparingly soluble in aqueous solution is a clue to their function. If released into the spermosphere and rhizosphere without a glycone group, the poor solubility will restrict transport to a small area close to the release site. Using a hydrophobic molecule as an inducer of

symbiosis is logical, because a molecule that diffuses too far from the appropriate location on the host would cause spurious and non-productive symbiotic elicitations (Shaw and Hooker, 2008). An effective chemoattractant, however, needs to form a robust, long-distance gradient for a cell to follow (Futrelle and Berg, 1972). This disparity in requirements makes finding a molecule that would effectively fulfill both roles problematic.

The odds of any individual bacterium seeding a nodule are minuscule (Denison and Kiers, 2011). The sheer density of all bacteria (which can approach  $10^8$  cells per gram of soil) around a plant's roots is in great excess to all potential sites of nodule formation (Torsvik et al., 1990; Roesch et al., 2007; Raynaud and Nunan, 2014). Most nodulation-competent rhizobia in the rhizosphere would exist elsewhere than at the tips of root hairs – the location of nodulation. As an alternative, the rhizobia that do not nodulate their host, along with the vast majority of other resident microbiota, can reasonably expect to survive on the exudates from the plant roots. It has been well documented that the rhizosphere is far richer in carbon than the surrounding bulk soil (Chiu et al., 2002; Guo-Mei Jia, 2015). Essentially all plants release root exudates, and the propensity to seek areas rich in carbon and nitrogen sources would allow rhizobia to acquire nutrients regardless of the source plant. This may be a superior survival strategy compared to seeking out a specific host plant for the express purpose of a low percentage chance of nodulation. Therefore, a flavonoid chemotaxis system would not be practical for the majority of circumstances rhizobia encounter, and since numerous other attractants are released from root hair tips, the bacterium would inherently swim to that location anyway, making a flavonoid sensing system redundant.

Our lab has two decades of experience in the study of multiple facets of rhizobial chemotaxis (Sourjik et al., 2000; Rotter et al., 2006; Meier et al., 2007; Meier and Scharf, 2009; Dogra et al., 2012; Shrestha et al., 2018). We have so far characterized the molecular sensing mechanisms of three different classes of chemoattractants. In particular, L-proline, sensed by McpU, has a chemotaxis coefficient of 100 (Webb et al., 2014). Stachydrine and other QACs sensed by McpX elicit chemotaxis coefficients around 80 (Webb et al., 2017a). McpV directly senses small monocarboxylates, but its ligands have chemotaxis coefficients of only 4 (Compton et al., 2018). While it is unwise to make direct comparisons between different strains, we note that Caetano-Anolles *et al.* reported a chemotaxis ratio of 2 to luteolin using their techniques and conditions (Caetanoanollés et al., 1988). Although this result agrees with our data on chemotaxis to flavonoids, we do not and cannot claim that taxis to flavonoids is distinguished from taxis to the

methanol cosolvent (Figs. 2.1 and 2.2). Given the magnitudinous differences in chemotaxis coefficients, we conclude that chemotaxis to amino acids and QACs, both of which are found in seed exudates, are the primary metabolites plant hosts can use to recruit *S. meliloti*. This conclusion is supported by chemotaxis experiments with synthetic mixtures of the amino acid spectrum in seed exudates, which showed that the amino acid fraction alone can elicit 23 % of the response to whole seed exudates (Webb et al., 2017b). In our hands, the signal to noise ratio of the capillary assay makes attractants with chemotaxis ratios below 3 difficult to identify with statistical significance. However, even if an attractant with such a low level of attraction could be accurately identified, its significance would still be dwarfed by far more potent attractants such as QACs and amino acids.

The paradigm of flavonoids acting as attractants for rhizobia was established several decades ago. This information became central to the thinking and models in the field of how the rhizobium-legume mutualism is initiated and established. A reexamination of past data in light of current knowledge on chemotaxis signaling in combination with the information we presented here, suggests that this paradigm is more of a plant-centric fallacy than a significant ecological phenomenon. Going forward, we hope the field acknowledges that the recruitment, culturing, and communication involved between bacteria and plant hosts is dependent on numerous chemical cues and diverse molecular mechanisms.

#### Acknowledgements

This study was supported by NSF grants MCB-1253234 and MCB-1817652 to Birgit Scharf. We are indebted to Benjamin Webb for preliminary explorations of the project, as well as members of the Scharf lab for critical review of the manuscript. The Virginia Tech Mass Spectrometry Incubator is partially maintained with funding from the Fralin Life Science Institute of Virginia Tech as well as NIFA (Hatch Accession Number 1015346).

## REFERENCES

1. Hazelbauer GL, Falke JJ, Parkinson JS. 2008. Bacterial chemoreceptors: high-performance signaling in networked arrays. *Trends Biochem Sci* 33:9-19.
2. Parkinson JS, Hazelbauer GL, Falke JJ. 2015. Signaling and sensory adaptation in *Escherichia coli* chemoreceptors: 2015 update. *Trends Microbiol* 23:257-66.
3. Bi S, Lai L. 2015. Bacterial chemoreceptors and chemoeffectors. *Cell Mol Life Sci* 72:691-708.
4. Krell T, Lacal J, Munoz-Martinez F, Reyes-Darias JA, Cadirci BH, Garcia-Fontana C, Ramos JL. 2011. Diversity at its best: bacterial taxis. *Environ Microbiol* 13:1115-24.
5. Miller LD, Russell MH, Alexandre G. 2009. Diversity in bacterial chemotactic responses and niche adaptation. *Adv Appl Microbiol* 66:53-75.
6. Lacal J, Garcia-Fontana C, Munoz-Martinez F, Ramos JL, Krell T. 2010. Sensing of environmental signals: classification of chemoreceptors according to the size of their ligand binding regions. *Environ Microbiol* 12:2873-84.
7. Scharf BE, Hynes MF, Alexandre GM. 2016. Chemotaxis signaling systems in model beneficial plant-bacteria associations. *Plant Molecular Biology* 90:549-559.
8. Brewin NJ. 1991. Development of the legume root nodule. *Annu Rev Cell Biol* 7:191-226.
9. Kondorosi E, Mergaert P, Kereszt A. 2013. A paradigm for endosymbiotic life: cell differentiation of *Rhizobium* bacteria provoked by host plant factors. *Annu Rev Microbiol* 67:611-28.
10. Haag AF, Arnold MF, Myka KK, Kerscher B, Dall'Angelo S, Zanda M, Mergaert P, Ferguson GP. 2013. Molecular insights into bacteroid development during *Rhizobium*-legume symbiosis. *FEMS Microbiol Rev* 37:364-83.
11. Drogue B, Dore H, Borland S, Wisniewski-Dye F, Prigent-Combaret C. 2012. Which specificity in cooperation between phytostimulating rhizobacteria and plants? *Res Microbiol* 163:500-10.
12. van Rhijn P, Vanderleyden J. 1995. The *Rhizobium*-plant symbiosis. *Microbiol Rev* 59:124-42.
13. Services NAS. 2017. Crop Production 2016 Summary. United States Department of Agriculture.
14. Lindeman WC, Glover CR, Flynn R, Idowu J. June, 2015 2015. Nitrogen Fixation by Legumes, on New Mexico State University Cooperative Extension. [http://aces.nmsu.edu/pubs/\\_a/A129/](http://aces.nmsu.edu/pubs/_a/A129/). Accessed 8/29/2017.
15. Webb BA, Helm RF, Scharf BE. 2016. Contribution of individual chemoreceptors to *Sinorhizobium meliloti* chemotaxis towards amino acids of host and nonhost seed exudates. *Mol Plant Microbe Interact* 29:231-239.

16. Webb BA, Hildreth S, Helm RF, Scharf BE. 2014. *Sinorhizobium meliloti* chemoreceptor McpU mediates chemotaxis toward host plant exudates through direct proline sensing. *Appl Environ Microbiol* 80:3404-15.
17. Webb BA, Compton KK, Castaneda Saldana R, Arapov T, Ray WK, Helm RF, Scharf BE. 2017. *Sinorhizobium meliloti* chemotaxis to quaternary ammonium compounds is mediated by the chemoreceptor McpX. *Mol Microbiol* 103:333-346.
18. Bernabeu-Roda L, Calatrava-Morales N, Cuellar V, Soto MJ. 2015. Characterization of surface motility in *Sinorhizobium meliloti*: regulation and role in symbiosis. *Symbiosis* 67:79-90.
19. Miller LD, Yost CK, Hynes MF, Alexandre G. 2007. The major chemotaxis gene cluster of *Rhizobium leguminosarum* bv. *viciae* is essential for competitive nodulation. *Mol Microbiol* 63:348-62.
20. Moe LA. 2013. Amino acids in the rhizosphere: from plants to microbes. *Am J Bot* 100:1692-705.
21. Barbour WM, Hattermann DR, Stacey G. 1991. Chemotaxis of *Bradyrhizobium japonicum* to soybean exudates. *Appl Environ Microbiol* 57:2635-9.
22. Nelson EB. 2004. Microbial dynamics and interactions in the spermosphere. *Annu Rev Phytopathol* 42:271-309.
23. Meier VM, Muschler P, Scharf BE. 2007. Functional analysis of nine putative chemoreceptor proteins in *Sinorhizobium meliloti*. *J Bacteriol* 189:1816-1826.
24. Meier VM, Scharf BE. 2009. Cellular localization of predicted transmembrane and soluble chemoreceptors in *Sinorhizobium meliloti*. *J Bacteriol* 191:5724-33.
25. Webb BA, Compton KK, Del Campo JSM, Taylor D, Sobrado P, Scharf BE. 2017. *Sinorhizobium meliloti* chemotaxis to multiple amino acids is mediated by the chemoreceptor McpU. *Mol Plant Microbe Interact* 30:770-777.
26. Ortega A, Zhulin IB, Krell T. 2017. Sensory repertoire of bacterial chemoreceptors. *Microbiol Mol Biol Rev* 81.
27. Anantharaman V, Aravind L. 2000. Cache - a signaling domain common to animal Ca(2<sup>+</sup>)-channel subunits and a class of prokaryotic chemotaxis receptors. *Trends Biochem Sci* 25:535-537.
28. Upadhyay AA, Fleetwood AD, Adebali O, Finn RD, Zhulin IB. 2016. Cache domains that are homologous to, but different from PAS domains comprise the largest superfamily of extracellular sensors in prokaryotes. *PLoS Comput Biol* 12:e1004862.
29. Garcia V, Reyes-Darias JA, Martin-Mora D, Morel B, Matilla MA, Krell T. 2015. Identification of a chemoreceptor for C2 and C3 carboxylic acids. *Appl Environ Microbiol* 81:5449-57.
30. Brewster JL, McKellar JL, Finn TJ, Newman J, Peat TS, Gerth ML. 2016. Structural basis for ligand recognition by a Cache chemosensory domain that mediates carboxylate sensing in *Pseudomonas syringae*. *Sci Rep* 6:35198.

31. Finn RD, Coghill P, Eberhardt RY, Eddy SR, Mistry J, Mitchell AL, Potter SC, Punta M, Qureshi M, Sangrador-Vegas A, Salazar GA, Tate J, Bateman A. 2016. The Pfam protein families database: towards a more sustainable future. *Nucleic Acids Research* 44:D279-D285.
32. Pokkuluri PR, Dwulit-Smith J, Duke NE, Wilton R, Mack JC, Bearden J, Rakowski E, Babnigg G, Szurmant H, Joachimiak A, Schiffer M. 2013. Analysis of periplasmic sensor domains from *Anaeromyxobacter dehalogenans* 2CP-C: Structure of one sensor domain from a histidine kinase and another from a chemotaxis protein. *Microbiologyopen* 2:766-777.
33. McKellar JLO, Minnell JJ, Gerth ML. 2015. A high-throughput screen for ligand binding reveals the specificities of three amino acid chemoreceptors from *Pseudomonas syringae* pv. *actinidiae*. *Molecular Microbiology* 96:694-707.
34. Adler J. 1973. A method for measuring chemotaxis and use of the method to determine optimum conditions for chemotaxis by *Escherichia coli*. *J Gen Microbiol* 74:77-91.
35. Owen AG, Jones DL. 2001. Competition for amino acids between wheat roots and rhizosphere microorganisms and the role of amino acids in plant N acquisition. *Soil Biology & Biochemistry* 33:651-657.
36. Jones DL, Darrah PR. 1994. Amino-Acid Influx at the Soil-Root Interface of Zea-Mays L and Its Implications in the Rhizosphere. *Plant and Soil* 163:1-12.
37. Jones DL, Edwards AC, Donachie K, Darrah PR. 1994. Role of Proteinaceous Amino-Acids Released in Root Exudates in Nutrient Acquisition from the Rhizosphere. *Plant and Soil* 158:183-192.
38. Odunfa VSA. 1979. Free Amino-Acids in the Seed and Root Exudates in Relation to the Nitrogen Requirements of Rhizosphere Soil *Fusaria*. *Plant and Soil* 52:491-499.
39. Futrelle RP, Berg HC. 1972. Specification of gradients used for studies of chemotaxis. *Nature* 239:517-8.
40. Wiseman T, Williston S, Brandts JF, Lin LN. 1989. Rapid measurement of binding constants and heats of binding using a new titration calorimeter. *Analytical Biochemistry* 179:131-137.
41. Reyes-Darias JA, Yang YL, Sourjik V, Krell T. 2015. Correlation between signal input and output in PctA and PctB amino acid chemoreceptor of *Pseudomonas aeruginosa*. *Molecular Microbiology* 96:513-525.
42. Lacal J, Alfonso C, Liu XX, Parales RE, Morel B, Conejero-Lara F, Rivas G, Duque E, Ramos JL, Krell T. 2010. Identification of a Chemoreceptor for Tricarboxylic Acid Cycle Intermediates. *Journal of Biological Chemistry* 285:23124-23134.
43. Fernandez M, Matilla MA, Ortega A, Krell T. 2017. Metabolic Value Chemoattractants Are Preferentially Recognized at Broad Ligand Range Chemoreceptor of *Pseudomonas putida* KT2440. *Frontiers in Microbiology* 8.
44. Ames P, Bergman K. 1981. Competitive Advantage Provided by Bacterial Motility in the Formation of Nodules by *Rhizobium-meliloti*. *Journal of Bacteriology* 148:728-729.

45. Gulash M, Ames P, Larosiliere RC, Bergman K. 1984. Rhizobia Are Attracted to Localized Sites on Legume Roots. *Applied and Environmental Microbiology* 48:149-152.
46. Garcia JAL, Barbas C, Probanza A, Barrientos ML, Manero FJG. 2001. Low molecular weight organic acids and fatty acids in root exudates of two *Lupinus* cultivars at flowering and fruiting stages. *Phytochemical Analysis* 12:305-311.
47. Zatakia HM, Arapov TD, Meier VM, Scharf BE. 2018. Cellular Stoichiometry of Methyl-Accepting Chemotaxis Proteins in *Sinorhizobium meliloti*. *Journal of Bacteriology* 200.
48. Cieslinski G, Van Rees KCJ, Szmigielska AM, Krishnamurti GSR, Huang PM. 1998. Low-molecular-weight organic acids in rhizosphere soils of durum wheat and their effect on cadmium bioaccumulation. *Plant and Soil* 203:109-117.
49. van Hees PAW, Dahlen J, Lundstrom US, Boren H, Allard B. 1999. Determination of low molecular weight organic acids in soil solution by HPLC. *Talanta* 48:173-179.
50. Li XL, Chen XM, Liu X, Zhou LC, Yang XQ. 2012. Characterization of soil low-molecular-weight organic acids in the Karst rocky desertification region of Guizhou Province, China. *Frontiers of Environmental Science & Engineering* 6:195-203.
51. Charles TC, Cai GQ, Aneja P. 1997. Megaplasmid and chromosomal loci for the PHB degradation pathway in *Rhizobium (Sinorhizobium) meliloti*. *Genetics* 146:1211-1220.
52. Dunn MF. 1998. Tricarboxylic acid cycle and anaplerotic enzymes in rhizobia. *Fems Microbiology Reviews* 22:105-123.
53. Biondi EG, Tatti E, Comparini D, Giuntini E, Mocali S, Giovannetti L, Bazzicalupo M, Mengoni A, Viti C. 2009. Metabolic Capacity of *Sinorhizobium (Ensifer) meliloti* Strains as Determined by Phenotype MicroArray Analysis. *Applied and Environmental Microbiology* 75:5396-5404.
54. Pickering BS, Oresnik IJ. 2008. Formate-dependent autotrophic growth in *Sinorhizobium meliloti*. *Journal of Bacteriology* 190:6409-6418.
55. Pellicer MT, Badia J, Aguilar J, Baldoma L. 1996. glc locus of *Escherichia coli*: Characterization of genes encoding the subunits of glycolate oxidase and the glc regulator protein. *Journal of Bacteriology* 178:2051-2059.
56. Lord JM. 1972. Glycolate Oxidoreductase in *Escherichia-Coli*. *Biochimica Et Biophysica Acta* 267:227-&.
57. Capela D, Barloy-Hubler F, Gouzy J, Bothe G, Ampe F, Batut J, Boistard P, Becker A, Boutry M, Cadieu E, Dreano S, Gloux S, Godrie T, Goffeau A, Kahn D, Kiss E, Lelaure V, Masuy D, Pohl T, Portetelle D, Puhler A, Purnelle B, Ramsperger U, Renard C, Thebault P, Vandenbol M, Weidner S, Galibert F. 2001. Analysis of the chromosome sequence of the legume symbiont *Sinorhizobium meliloti* strain 1021. *Proceedings of the National Academy of Sciences of the United States of America* 98:9877-9882.
58. Wibberg D, Blom J, Ruckert C, Winkler A, Albersmeier A, Puhler A, Schluter A, Scharf BE. 2013. Draft genome sequence of *Sinorhizobium meliloti* RU11/001, a model organism for flagellum structure, motility and chemotaxis. *J Biotechnol* 168:731-3.



59. Schweinitzer T, Josenhans C. 2010. Bacterial energy taxis: a global strategy? Archives of Microbiology 192:507-520.
60. Hou SB, Larsen RW, Boudko D, Riley CW, Karatan E, Zimmer M, Ordal GW, Alam M. 2000. Myoglobin-like aerotaxis transducers in Archaea and Bacteria. Nature 403:540-544.
61. Glekas GD, Mulhern BJ, Kroc A, Duelfer KA, Lei V, Rao CV, Ordal GW. 2012. The Bacillus subtilis Chemoreceptor McpC Senses Multiple Ligands Using Two Discrete Mechanisms. Journal of Biological Chemistry 287.
62. Richarme G. 1982. Interaction of the maltose-binding protein with membrane vesicles of Escherichia coli. J Bacteriol 149:662-7.
63. Koiwai O, Hayashi H. 1979. Studies on bacterial chemotaxis. IV. Interaction of maltose receptor with a membrane-bound chemosensing component. J Biochem 86:27-34.
64. Bertani G. 1951. Studies on Lysogenesis .1. The Mode of Phage Liberation by Lysogenic Escherichia-Coli. Journal of Bacteriology 62:293-300.
65. Götz R, Limmer N, Ober K, Schmitt R. 1982. Motility and Chemotaxis in 2 Strains of Rhizobium with Complex Flagella. Journal of General Microbiology 128:789-798.
66. Sourjik V, Schmitt R. 1996. Different roles of CheY1 and CheY2 in the chemotaxis of Rhizobium meliloti. Molecular Microbiology 22:427-436.
67. Bryksin AV, Matsumura I. 2010. Overlap extension PCR cloning: a simple and reliable way to create recombinant plasmids. Biotechniques 48:463-465.
68. Simon R, O'Connell M, Labes M, Pühler A. 1986. Plasmid vectors for the genetic analysis and manipulation of rhizobia and other gram-negative bacteria. Methods Enzymol 118:640-659.
69. Simon R, Priefer U, Pühler A. 1983. A broad host range mobilisation system for *in vivo* genetic engineering: Transposon mutagenesis in gram negative bacteria. Bio/Technology 1:783-791.
70. Biasini M, Bienert S, Waterhouse A, Arnold K, Studer G, Schmidt T, Kiefer F, Gallo Cassarino T, Bertoni M, Bordoli L, Schwede T. 2014. SWISS-MODEL: modelling protein tertiary and quaternary structure using evolutionary information. Nucleic Acids Res 42:W252-8.
71. Arnold K, Bordoli L, Kopp J, Schwede T. 2006. The SWISS-MODEL workspace: a web-based environment for protein structure homology modelling. Bioinformatics 22:195-201.
72. Peitsch MC, Schwede T, Guex N. 2000. Automated protein modelling--the proteome in 3D. Pharmacogenomics 1:257-66.
73. Kiefer F, Arnold K, Kunzli M, Bordoli L, Schwede T. 2009. The SWISS-MODEL Repository and associated resources. Nucleic Acids Res 37:D387-92.
74. Biasini M, Bienert S, Waterhouse A, Arnold K, Studer G, Schmidt T, Kiefer F, Cassarino TG, Bertoni M, Bordoli L, Schwede T. 2014. SWISS-MODEL: modelling protein tertiary and quaternary structure using evolutionary information. Nucleic Acids Research 42:W252-W258.

75. Guex N, Peitsch MC, Schwede T. 2009. Automated comparative protein structure modeling with SWISS-MODEL and Swiss-PdbViewer: A historical perspective. *Electrophoresis* 30:S162-S173.
76. Kiefer F, Arnold K, Kunzli M, Bordoli L, Schwede T. 2009. The SWISS-MODEL Repository and associated resources. *Nucleic Acids Research* 37:D387-D392.
77. Gasteiger E, Hoogland C, Gattiker A, Duvaud S, Wilkins MR, Appel RD, Bairoch A. 2005. Protein identification and analysis tools on the ExPASy server. In Walker JM (ed), *The proteomics handbook*, Humana Press, Totowa, NJ:571-607.
78. Hanahan D, Meselson M. 1983. Plasmid screening at high colony density. *Methods Enzymol* 100:333-42.
79. Pleier E, Schmitt R. 1991. Expression of two *Rhizobium meliloti* flagellin genes and their contribution to the complex filament structure. *J Bacteriol* 173:2077-2085.
80. Schäfer A, Tauch A, Jager W, Kalinowski J, Thierbach G, Pühler A. 1994. Small mobilizable multi-purpose cloning vectors derived from the *Escherichia coli* plasmids pK18 and pK19: selection of defined deletions in the chromosome of *Corynebacterium glutamicum*. *Gene* 145:69-73.

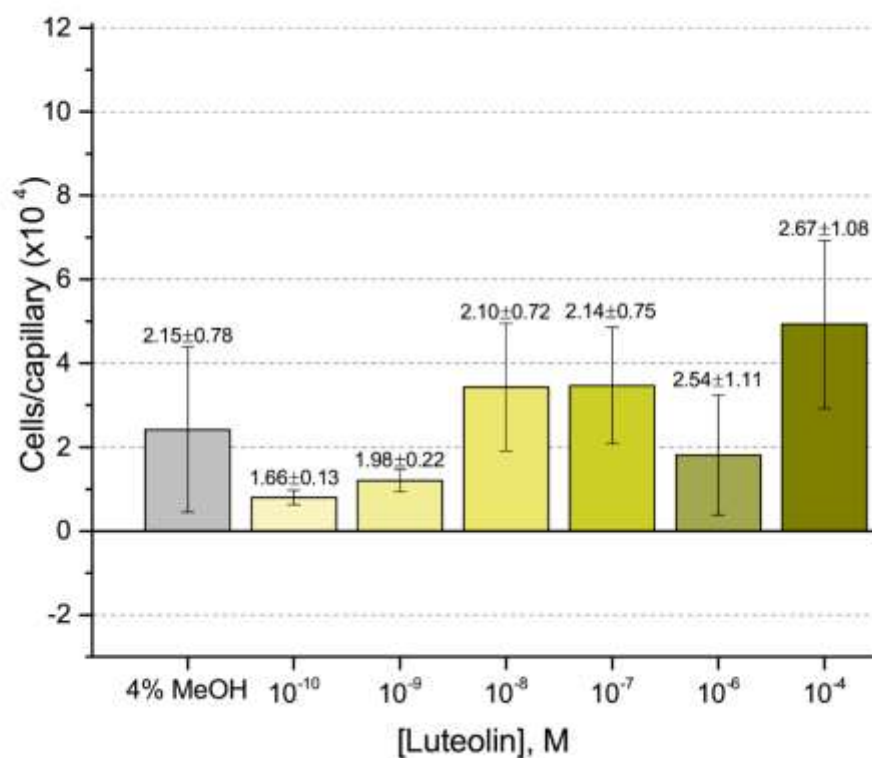
**Table 2.1. Levels of Selected Flavonoids from Seed Exudates**

Five flavonoids were quantified in alfalfa seed exudates using UPLC-MS in MRM, positive ion mode. Seed exudates were harvested by incubating surface-sterilized alfalfa seeds in water for 24 hours and enriched for flavonoids by solid phase extraction. Values are the means and standard deviations from six independent replicates.

<b>Flavonoid</b>	<b>MRM Transition</b>	<b>ng/seed</b>	<b>pmol/seed</b>
Chrysoeriol <sup>†</sup>	301.05>286.00	5 ± 1	16 ± 4
Hyperoside	465.00>303.05	2,455 ± 32	4,963 ± 66
Luteolin <sup>†</sup>	287.15>153.05	38 ± 9	133 ± 31
Luteolin-7-glucoside	449.10>287.00	52 ± 43	117 ± 96
Quercetin*	303.15>153.15	< LOQ	< LOQ

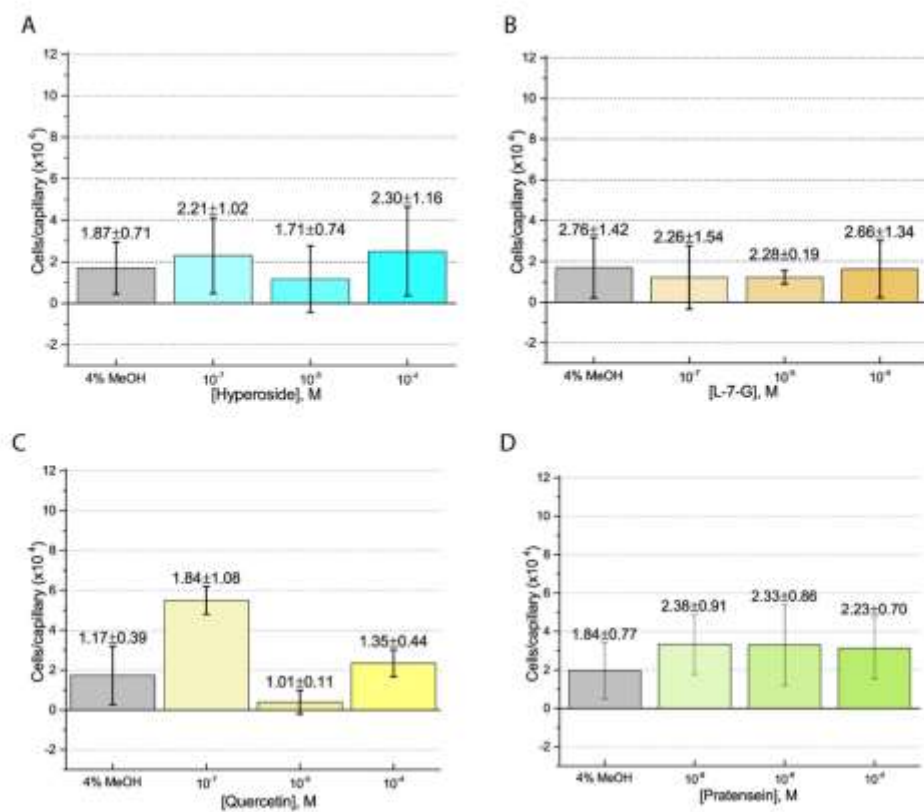
<sup>†</sup> Indicates nod gene-inducing flavonoid in *S. meliloti*.

\*LOQ, Limit of Quantification (0.5 ng/seed).



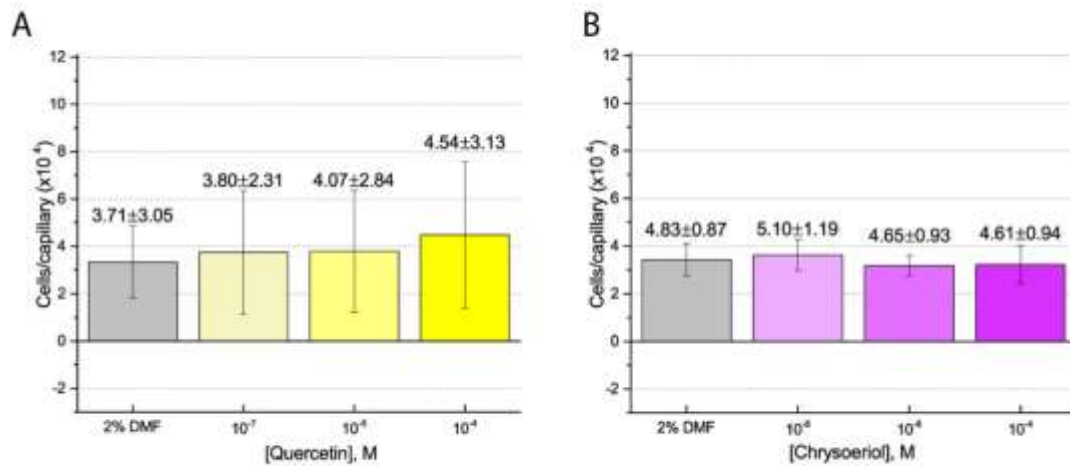
**Fig 2.1. Capillary chemotaxis assay of *S. meliloti* to luteolin and methanol**

Each concentration of luteolin tested included 4 % methanol in the attractant solution. The bars are the means and standard deviation of three biological replicates in which the number of cells that accumulated in a reference capillary was subtracted from that of the test capillary. The numbers above the bars represent the means and standard deviations of the chemotaxis ratio where the number of cells that accumulated in the test capillary were divided by the number of cells that accumulated in a reference capillary.

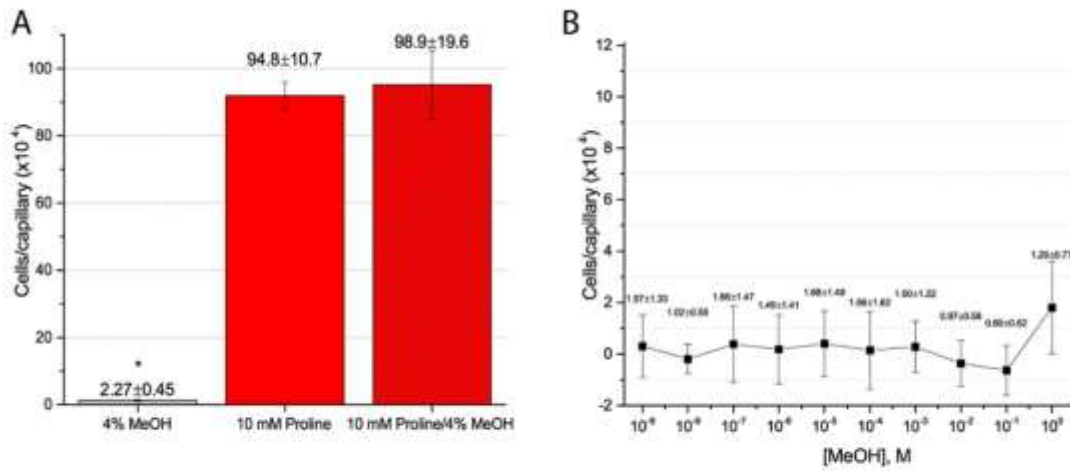


**Fig 2.2. Capillary chemotaxis assays of *S. meliloti* to flavonoids in methanol**

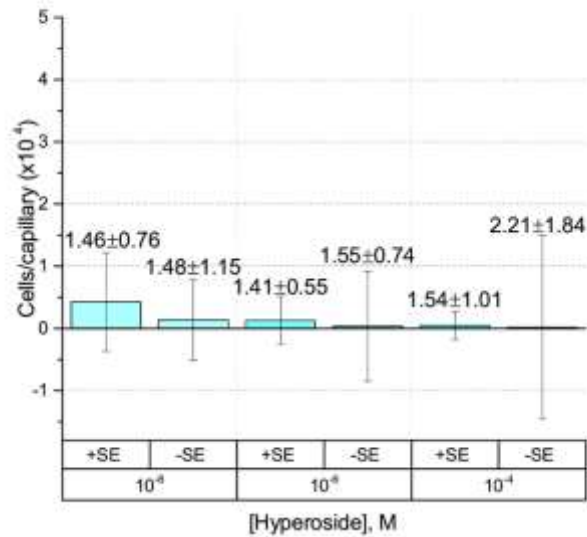
Each concentration of flavonoid tested included 4 % methanol in the attractant solution. The bars are the means and standard deviation of three biological replicates in which the number of cells that accumulated in a reference capillary was subtracted from that of the test capillary. The numbers above the bars represent the means and standard deviations of the chemotaxis ratio where the number of cells that accumulated in the test capillary were divided by the number of cells that accumulated in a reference capillary. (A). Hyperoside; (B). Luteolin-7-glucoside; (C). Quercetin; (D). Pratensein.



**Fig 2.3. Capillary chemotaxis assays of *S. meliloti* to flavonoids in dimethylformamide (DMF)**  
 Each concentration of flavonoid tested included 2 % DMF in the attractant solution. Values are the means and standard deviations of three biological replicates in which the number of cells that accumulated in a reference capillary was subtracted from that of the test capillary. The numbers above the bars represent the means and standard deviations of the chemotaxis ratio where the number of cells that accumulated in the test capillary were divided by the number of cells that accumulated in a reference capillary. **(A)**. Quercetin; **(B)**. Chrysoeriol.



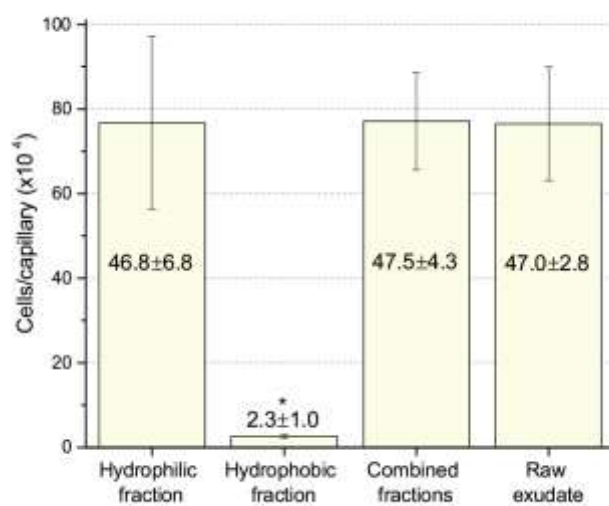
**Fig 2.4. Capillary chemotaxis assays testing the effect of methanol on *S. meliloti* chemotaxis.** (A). Comparison of chemotaxis to 4 % methanol, 10 mM proline, and 10 mM proline + 4 % methanol. The bars are the means and standard deviation of three biological replicates in which the number of cells that accumulated in a reference capillary was subtracted from that of the test capillary. The asterisk denotes  $P < 0.0008$  using Student's t-test (B). Dose response curve to methanol. The means and standard deviations were calculated with four biological replicates performed in technical duplicate. The numbers above the bars represent the mean and standard deviation of the chemotaxis ratio where the number of cells that accumulated in the test capillary were divided by the number of cells that accumulated in a reference capillary. Note the difference in scale between (A) and (B).



**Fig 2.5. Capillary chemotaxis assays of *S. meliloti* to hyperoside in the absence of cosolvent**

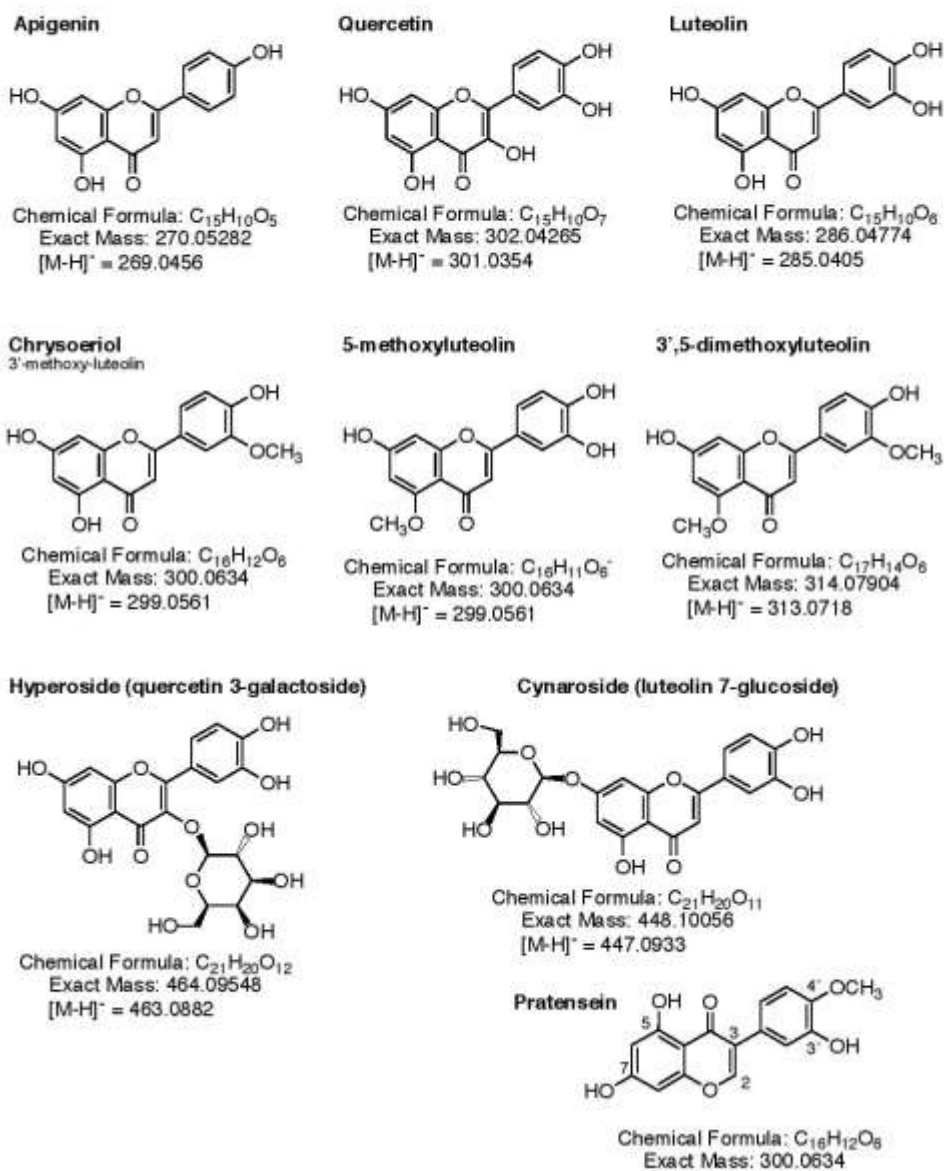
Data from cultures grown in the presence of 5 % seed exudates (SE) are indicated with the label +SE on the x-axis. Cultures grown without seed exudates are indicated with the label -SE. The bars are the means and standard deviation of three biological replicates in which the number of cells that accumulated in a reference capillary was subtracted from that of the test capillary. The numbers above the bars represent the means and standard deviations of the chemotaxis ratio where the number of cells that accumulated in the test capillary were divided by the number of cells that accumulated in a reference capillary. Note that the of the y-axis scale is 50 % of all preceding experiments.



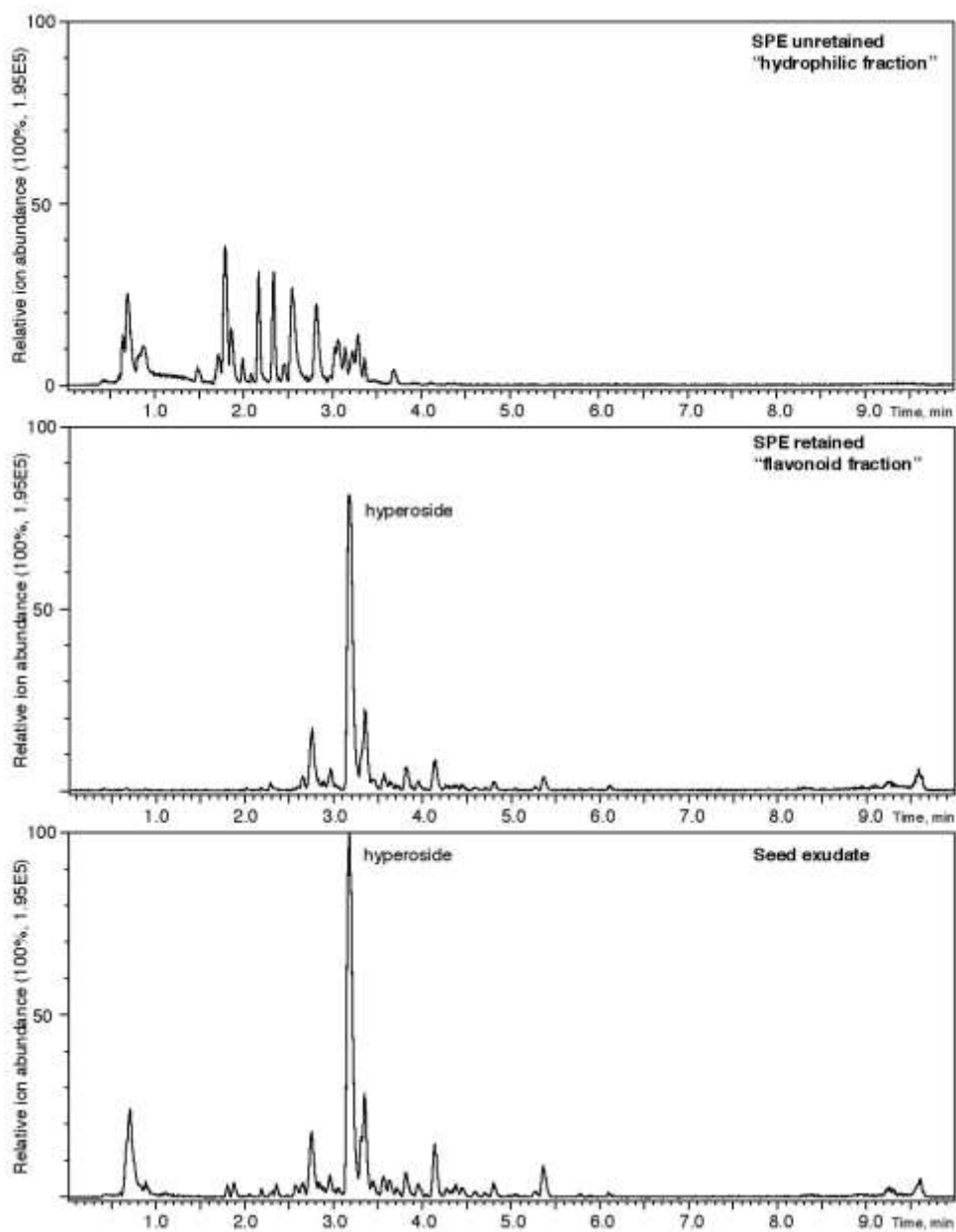


**Fig 2.6. Chemotaxis of *S. meliloti* to fractionated and raw alfalfa seed exudates**

Comparison of the chemotactic potential of seed exudate fractions. The hydrophilic fraction is the combined flow-through and water wash. The hydrophobic fraction is the combined methanol washes. Combined fractions is the mixture of hydrophilic and hydrophobic fractions in equal proportion. Raw exudates were not separated by SPE. All fractions were used at 0.8-fold of their original concentration. The asterisk denotes  $P < 0.011$  using Student's t-test.

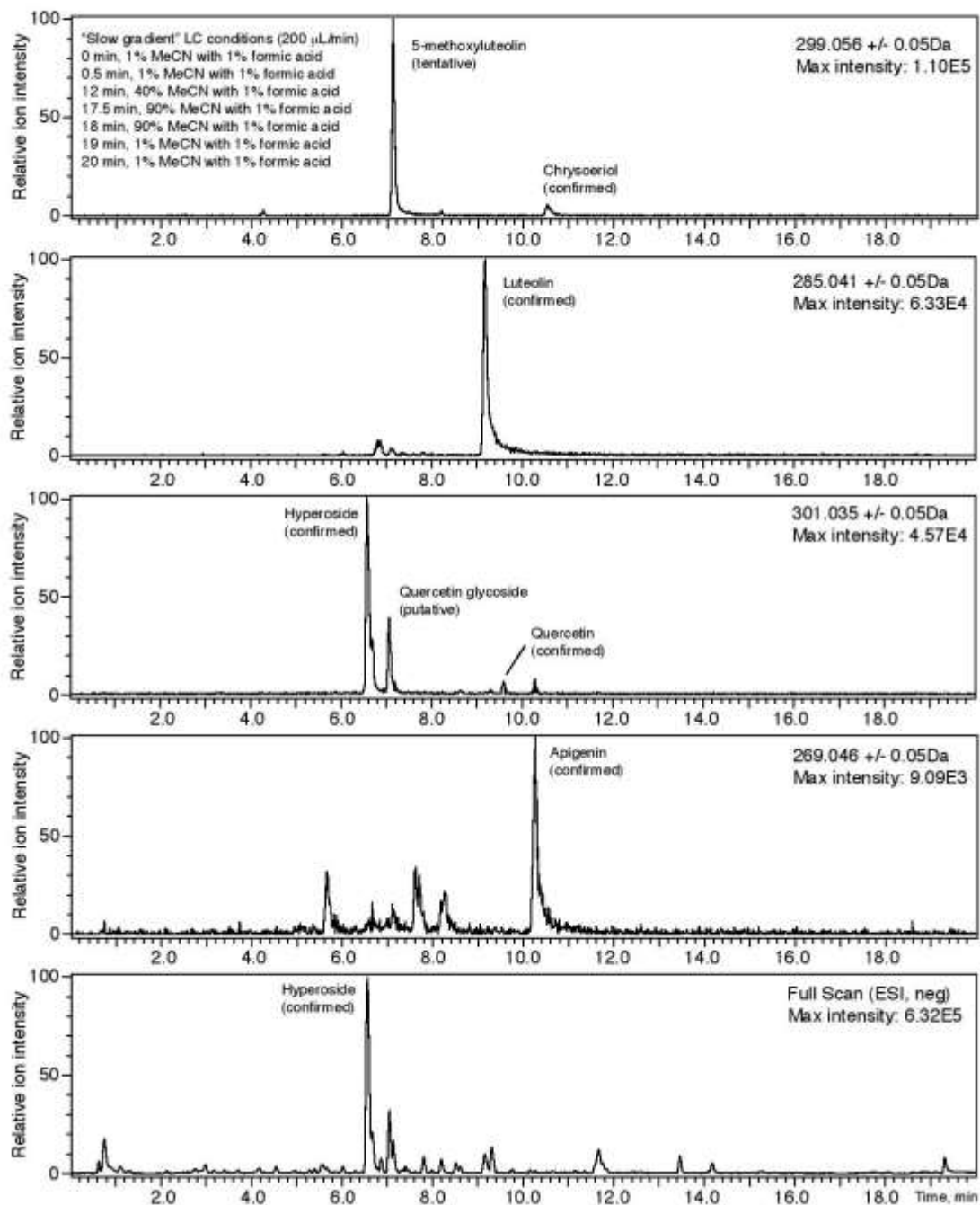


**Fig. 2.7. Structures of flavonoids.**

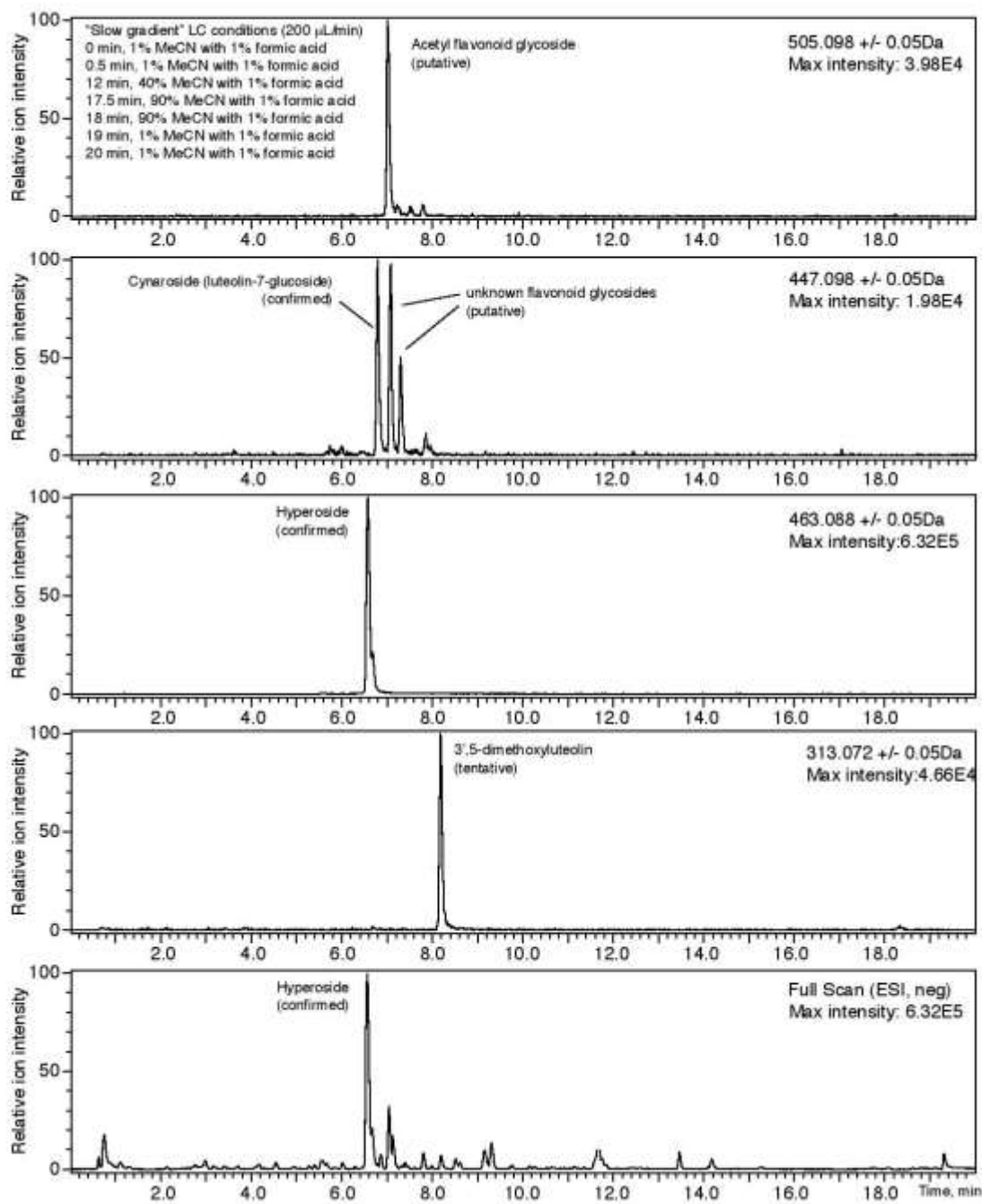


**Fig. 2.8. LC-MS chromatogram in base peak mode**

Traces are normalized to relative ion abundance of hyperoside in the original seed exudates. “Fast gradient” LC conditions. Hyperoside was confirmed with an authentic standard (retention time, parent mass, and fragmentation pattern).



**Fig. 2.9. Selected ion monitoring for luteolin-, quercetin- and apigenin-based compounds**  
 Confirmed: authentic standard, retention time, parent mass, and fragmentation pattern matches.  
 Tentative: previous literature, no authentic standard. Putative: parent mass and fragmentation pattern only.



**Fig. 2.10. Selected ion monitoring for luteolin- and quercetin-based compounds**

Confirmed: authentic standard, retention time, parent mass, and fragmentation pattern matches.

Tentative: previous literature, no authentic standard. Putative: parent mass and fragmentation pattern only.

**Chapter 3 - *Sinorhizobium meliloti* chemoreceptor McpV senses short chain carboxylates  
via direct binding**

K. KARL COMPTON<sup>a</sup>, SHERRY B. HILDRETH<sup>a</sup>, RICHARD F. HELM<sup>b</sup>, BIRGIT E.  
SCHARF<sup>a#</sup>

<sup>a</sup> Virginia Tech, Department of Biological Sciences, Life Sciences I, Blacksburg, VA 24061

<sup>b</sup> Virginia Tech, Department of Biochemistry, VA 24061

Running title: McpV is a sensor for small carboxylates

Key words: chemotaxis, plant host exudate, motility, rhizosphere, symbiosis

# For correspondence:

E-mail bscharf@vt.edu

Tel (+1) 540 231 0757

Fax (+1) 540 231 4043

Biological Sciences, Life Sciences I

Virginia Tech

Blacksburg, VA 24061, USA

**Journal of Bacteriology. J Bacteriol 200:e00519-18. Accepted manuscript posted online 10<sup>th</sup>  
of September, 2018. <https://doi.org/10.1128/JB.00519-18>.**

**Attribution: KKC designed and carried out experiments except carboxylate analysis and  
quantification. BES provided experimental guidance and manuscript editing. SBH  
analyzed and quantified carboxylates in exudates. RFH provided oversight on  
metabolomics**

## ABSTRACT

*Sinorhizobium meliloti* is a soil-dwelling endosymbiont of alfalfa with eight chemoreceptors to sense environmental stimuli during its free-living state. The functions of two receptors have been characterized, with McpU and McpX serving as general amino acid and quaternary ammonium compound sensors, respectively. Both receptors use a dual Cache domain for ligand binding. We identified that the ligand-binding, periplasmic region (PR) of McpV contains a single Cache domain. Homology modeling revealed that McpV<sup>PR</sup> is structurally similar to a sensor domain of a chemoreceptor with unknown function from *Anaeromyxobacter dehalogenans*, which crystallized with acetate in its binding pocket. We therefore assayed McpV for carboxylate binding and *S. meliloti* for carboxylate sensing. Differential scanning fluorimetry identified ten potential ligands for McpV<sup>PR</sup>. Nine of these are monocarboxylates with chain lengths between two and four carbons. We selected seven compounds for capillary assay analysis, which established positive chemotaxis of *S. meliloti* wild type with concentrations of peak attraction at 1 mM for acetate, propionate, pyruvate, and glycolate, and 100 mM for formate and acetoacetate. Deletion of *mcpV* or mutation of residues essential for ligand coordination abolished positive chemotaxis to carboxylates. Using microcalorimetry we determined that dissociation constants of the seven ligands with McpV<sup>PR</sup> were in the micromolar range. An McpV<sup>PR</sup> variant with a mutation in the ligand coordination site displayed no binding to isobutyrate or propionate. Of all the carboxylates tested as attractants, only glycolate was detected in alfalfa seed exudates. This work examines the relevance of carboxylates and their sensor to the rhizobium-legume interaction.

## IMPORTANCE

Legumes share a unique association with certain soil-dwelling bacteria known broadly as the rhizobia. Through concerted interorganismal communication, a legume allows the intracellular infection by its cognate rhizobial species. The plant then forms an organ, the root nodule, dedicated to housing and supplying fixed carbon and nutrients to the bacteria. In return, the engulfed rhizobia, differentiated into bacteroids, fix atmospheric N<sub>2</sub> into ammonium for the plant host. This interplay is of great benefit to the cultivation of legumes, such as alfalfa and soybeans, and is initiated by chemotaxis to the host plant. This study on carboxylate chemotaxis contributes to the understanding of rhizobial survival and competition in the rhizosphere and aids the development of commercial inoculants.

## INTRODUCTION

Motility and navigation are two behaviors that bacteria exhibit to choose an optimal environment for survival and growth. Flagellar-driven motility is regulated by a finely tuned sensory array and a two-component signal transduction system that ultimately controls flagellar motor rotation. In *Escherichia coli*, the initial step in sensing is the binding of a ligand to its cognate methyl-accepting chemotaxis protein or MCP. Ligand binding typically occurs in the periplasmic region (PR) of the chemoreceptor, initiating a molecular stimulus that is transferred through the cytoplasmic membrane. Upon attractant binding autophosphorylation of the histidine kinase CheA is inhibited. Consequently, the corresponding response regulator, CheY, remains unphosphorylated and inactive, leading to unaltered flagellar motor rotation and a smooth swimming path of the cell. In the absence of ligand binding, CheA phosphorylates and activates CheY, which will bind to the flagellar motor and induce a tumble behavior. During tumbles, the bacterium can randomly reorient to a new direction. This behavior, called biased random walk, allows the bacterium to swim towards attractants and away from repellents (1-4).

The genomes of bacteria are expected to contain varying numbers and types of chemoreceptors that reflect their niche and lifestyle requirements. Denizens of static environments and simple niches are found to have few to no chemoreceptors, while those that share a complex interplay with other organisms or have diverse metabolic capabilities encode far more chemoreceptors in their genomes (4-6). The Alphaproteobacteria typify the latter case with organisms such as *Azospirillum lipoferum*, *Bradyrhizobium* sp. BTAil, and *Rhizobium phaseoli* containing 63, 60, and 29 predicted chemoreceptors, respectively (7). These organisms colonize the roots of plants and promote plant growth by fixing atmospheric nitrogen and outcompeting plant pathogens. Within the Alphaproteobacteria is the *Rhizobiaceae*, a bacterial family that forms a species-specific endosymbiosis with members of the *Fabaceae* plant family. The rhizobium and plant host communicate and undergo highly specific developmental changes that ultimately lead to the formation of a root organ called a nodule. Within these nodules differentiated bacteroids occupy membranous organelles inside host cells and the plant provides the bacteroids metabolizable carbon sources to fuel the fixation of nitrogen gas to ammonium (8-12).

*Sinorhizobium (Ensifer) meliloti* is the cognate symbiont for alfalfa (*Medicago sativa* L.), an important forage crop that the United States produced over 58 million tons in 2016 (13). Alfalfa and other legumes capable of nitrogen-fixing symbiosis can be grown largely free of costly and



environmentally deleterious nitrogenous fertilizers that may leach into neighboring ecosystems (14). Plants recruit *S. meliloti* and other soil microorganisms to the rhizosphere with the plethora of chemicals exuded from the roots. These compounds include amino acids, quaternary ammonium compounds, sugars, and organic acids, to name a few. While not directly involved in the symbiotic process, chemotaxis is critical to competition for root nodule occupancy (7, 15-22).

Nine putative chemoreceptors, namely McpS through McpZ, and the internal chemoreceptor, IcpA, are encoded in the genome of most *S. meliloti* strains. Previous studies have demonstrated that *mcpS* was not expressed when cells were motile and chemotactically active. Therefore, it was concluded that McpS is utilized in cellular processes other than chemotaxis (23, 24). The function of two of the eight chemoreceptors involved in *S. meliloti* chemotaxis have been elucidated. McpU is a general amino acid receptor, sensing all non-acidic proteogenic amino acids, as well as several non-proteogenic amino acids (16, 25). McpX senses quaternary ammonium compounds (QACs) such as glycine betaine, trigonelline, and choline through direct binding, and is the first QAC chemoreceptor described in bacteria (17). Amino acids and QACs are exuded by germinating alfalfa seeds in chemotactically relevant concentrations (15, 17). The PR of McpU and McpX both contain a dCache\_1 (dual calcium channels and chemotaxis receptors) domain. The interaction of Cache domains with small molecules is well described. A major fraction of extracellular sensors in prokaryotes employ Cache domains (16, 17, 26-31). Besides McpU and McpX, *S. meliloti* possesses a third Cache domain containing chemoreceptor, McpV, which has an sCache\_2 (single Cache) domain in its PR.

In this work, we screened the ligand profile of the purified McpV periplasmic region (McpV<sup>PR</sup>) and characterized ligand interaction in direct binding studies. Chemotaxis of *S. meliloti* wild type but not *mcpV* mutant strains to various carboxylates was established with traditional capillary assays, confirming the role of McpV as a carboxylate sensor. We hypothesized that carboxylate exudation by alfalfa recruits its symbiont to the rhizosphere, which was tested by quantifying these compounds in germinating alfalfa seed exudates. The knowledge accrued from this study establishes short chain carboxylates as another facet of the legume-rhizobium interplay and will inform future research on improving legume symbiosis for the benefit of agriculture.

## RESULTS

### **A structure-based homology search suggests interaction of McpV with acetate**

The periplasmic region of McpV (McpV<sup>PR</sup>) encompasses a conserved sCache (calcium channels and chemotaxis receptors) signaling domain (27) (amino acid residues 35-177; (23, 24)). A homology search in the SWISS-MODEL repository revealed that McpV<sup>PR</sup> shares sequence identity (53.6%) with the sensor domain of Adeh\_3718, an uncharacterized chemoreceptor from *Anaeromyxobacter dehalogenans*.(32). SWISS-Model generated a structural model of McpV<sup>PR</sup> using the PR of Adeh\_3718 (PDB entry 4K08; (32)). The global mean quality estimate is 0.75, indicating a high quality model. The PR of Adeh\_3718 was crystallized in complex with acetate, suggesting that McpV<sup>PR</sup> might also bind acetate. In the Adeh\_3718 structure, the carboxylate group of acetate forms salt bridges to the side chains of His107 and Lys160 and hydrogen bonds are found with the side chains of Tyr94 and Tyr147. In the homology model of McpV<sup>PR</sup>, these four ligand-coordinating residues are conserved with the corresponding residues in McpV being identified as His103, Lys156, Tyr90, Tyr143 (Fig. 3.1).

### **A high-throughput differential scanning fluorimetry assay screens the putative ligand profile of McpV**

The discovery of acetate in the binding pocket of the homology model suggests carboxylates as possible ligands for McpV. To investigate this possibility, a high-throughput *in-vitro* differential scanning fluorimetry (DSF) assay was used to screen the ligand profile of recombinantly expressed and purified McpV<sup>PR</sup>. Biolog plate PM1 was used for this screen because it contains a range of carbon sources such as sugars, carboxylates, nucleotides, detergents, and amino acids (16, 33). The  $T_m$  of the McpV<sup>PR</sup> in the presence of most compounds was within  $\pm 2^\circ\text{C}$  of the water control ( $57^\circ\text{C}$ ), therefore an interaction was defined as a  $T_m$  shift greater than  $3^\circ\text{C}$ . The screen identified ten compounds that interacted with McpV<sup>PR</sup> in monophasic melting reactions (Fig. 3.2, Table 3.1). With the exception of methyl pyruvate, all of these compounds are monocarboxylates with chain lengths between two and four carbons. With a  $\Delta T_m$  of  $12.3^\circ\text{C}$ , acetate and propionate elicited the greatest thermal shifts. Pyruvate caused the third greatest shift of  $11.8^\circ\text{C}$ , while glycolate and L-lactate shifted the  $T_m$  by  $9.3$  and  $8.2^\circ\text{C}$ , respectively. Acetoacetate, a four-carbon carboxylate, produced the next greatest shift of  $6.0^\circ\text{C}$ , followed by glyoxylate and methyl pyruvate, the latter being the only ester that caused a significant shift. Finally,  $\alpha$ -hydroxybutyrate and  $\alpha$ -ketobutyrate elicited the two lowest shifts of  $4.0$  and  $3.8^\circ\text{C}$ , respectively.

Since Tyr143 is one of the four key residues presumably involved in ligand coordination, we screened the purified McpV<sup>Y143A-PR</sup> variant for its interaction with small molecules using Biolog plate PM1. The most drastic change was the reduction of the melting temperature in the absence of a putative ligand from 57 to 42.5°C. The overall ranking of molecules by  $\Delta T_m$  was somewhat maintained (Table 1). Acetate, propionate, and pyruvate induced the three greatest  $\Delta T_m$  of 9.0, 9.0, and 9.5°C, respectively, as compared to the approximate 12.0°C shift for each with the wild-type protein. The shift caused by glycolate also decreased from 9.3 to 6.0°C. The shift produced by acetoacetate was 6°C for both the variant and wild-type protein. The shift elicited by L-lactate was severely reduced from 11.8°C with the wild-type protein to 2.0°C with the variant protein. Glyoxylate and methyl-pyruvate caused shifts of 4.5 and 4.0°C, respectively, reduced from approximately 6.0°C when compared to the wild-type protein. The only carboxylate to cause a greater shift in the substitution variant than in the wild-type protein was alpha-ketobutyrate, which increased from 3.8°C in the wild-type protein to 4.5°C in the mutant variant. Lastly, alpha-hydroxybutyrate produced an insignificant shift of 1.5°C with the mutant variant compared to 4°C with the wild-type protein (Table 1).

In conclusion, carboxylates with two to four carbons interact with McpV<sup>PR</sup>, the two- and three-carbon carboxylates causing a larger temperature shift than the four-carbon carboxylates. When Tyr143 is substituted for alanine, the stabilizing effect of many of the compounds tested was greatly reduced, implicating this residue in small molecule interaction.

### **Chemotaxis of *S. meliloti* wild type to carboxylates**

The ultimate reaction that results from ligand-chemoreceptor interaction is the translocation of the bacterium to the source of attractants or away from repellents. The traditional capillary assay allows for chemotactic responses to be quantified and classified (34). Formate, acetate, propionate, and butyrate were all tested to compare the simplest carboxylates of each chain length. Pyruvate, glycolate, and acetoacetate are of physiological relevance and were tested to compare the effects of their different functional groups. Each compound elicited a dose-dependent reaction curve from *S. meliloti* wild type (RU11/001) that peaked and subsequently declined, as is characteristic of an attractant chemotactic behavior (Fig. 3.3). *S. meliloti* was attracted to acetate, propionate, pyruvate, and glycolate with a peak attraction at 1 mM, with glycolate also recruiting nearly as many bacteria at 10 mM. The response curve to butyrate formed a broad plateau between 0.1 and 10 mM.

Attraction to formate peaked at 100 mM, but dropped to near zero at the two flanking concentrations tested. Attraction to acetoacetate was also highest at 100 mM, but its curve shared the profile of acetate and pyruvate, rather than that of formate. When comparing accumulation of cells, acetate, propionate, and acetoacetate were the most potent attractants, drawing around 110,000 cells to the capillary. Pyruvate and glycolate followed, with 85,000 and 74,000 cells, respectively. Formate and butyrate ranked last, accumulating only 35,000 to 36,000 cells per capillary on average (Fig. 3.3). To attribute the observed accumulation of bacteria to chemotaxis, a strain lacking all nine chemoreceptors, RU13/149, was tested in the chemotaxis assay at concentrations of peak attraction for four representative compounds, namely formate, acetate, propionate, and butyrate (Fig 3.4). As predicted, chemotaxis to each of the four compounds tested was completely abolished (Fig. 3.4). Together, these data demonstrate that one to four carbon carboxylates are chemoattractants for *S. meliloti*.

### **McpV is mediating carboxylate chemotaxis in *S. meliloti***

The DSF analysis identified McpV as a potential chemoreceptor for carboxylates. To assess the impact of *mcpV* on carboxylate chemotaxis, a strain lacking *mcpV*, RU11/830, was tested at concentrations of peak attraction for all seven carboxylates (Fig. 3.5A). In the absence of *mcpV*, chemotaxis to carboxylates was not detected. We next verified that the deletion of *mcpV* had no negative impact on chemotaxis in general. When proline chemotaxis was compared for wild type and the *mcpV* deletion strain, no reduction of proline attraction was observed. It should be noted that chemotaxis to 10 mM proline was improved by 1.5 fold in the absence of *mcpV*. For comparison to carboxylate taxis, chemotaxis of *S. meliloti* wild type to 10 mM L-proline drew about 460,000 bacteria to the capillary, which was more than four times that of any of the carboxylates (Fig. 3.5B). Therefore, carboxylates are less effective as attractants than proline.

The homology model of McpV<sup>PR</sup> revealed several conserved residues that appear to play a role in ligand binding. To test the role of two of these residues, *S. meliloti* strains harboring a Y143A (BS232) or H103E (BS234) substitution in McpV were constructed and tested in the capillary assay. Neither mutant strain exhibited any chemotaxis to 1 mM propionate, establishing a role of *mcpV* in carboxylate chemotaxis in *S. meliloti* (Fig. 3.6). In addition, residues Tyr143 and His103 are essential components of the McpV ligand binding pocket.

### **Isothermal titration calorimetry demonstrates direct binding of carboxylates to McpV<sup>PR</sup>**

To validate that carboxylate chemotaxis in *S. meliloti* is mediated through direct binding to McpV and to determine binding parameters, we performed isothermal titration calorimetry (ITC) at 25°C (except for acetoacetate, which was titrated at 28°C). All compounds displayed binding through the generation of exothermic binding reactions. Data were fitted using the “one binding sites” model and dissociation constants ( $K_d$ ) were calculated. Propionate and acetate exhibited the tightest binding with a  $K_d$  of 3.4 and 9.1  $\mu\text{M}$ , respectively (Figs. 3.7B, 3.7C). The next tightest binding occurred for glycolate and pyruvate with a  $K_d$  of 27 and 33  $\mu\text{M}$  each, followed by acetoacetate with a  $K_d$  of 280  $\mu\text{M}$  (Figs. 3.7E, 3.7F, 3.7G). Lastly, formate had the lowest affinity with a  $K_d$  of approximately 8.7 mM (Fig. 3.7A). Butyrate and isobutyrate titrations resulted in an exothermic isotherm that quickly transitioned into endothermic reactions (Figs. 3.7D, 3.7H). For isobutyrate, this pattern of interaction occurred at 28, 25 and 15°C. Dissociation constants were not determined for these two compounds because the shape of the curve could not be fit appropriately with the “one binding sites”.. To confirm the capillary assay results obtained for strains with mutations in the McpV binding pocket, purified McpV<sup>Y143A-PR</sup> at a concentration of 75  $\mu\text{M}$  was titrated against 15 mM propionate. The dissociation constant for propionate was approximately 2.5 mM, 1,000 fold lower compared to the wild type (Figs. 3.7C, 3.8). Together, the ITC data validated results gained from DSF experiments, established direct binding of carboxylates to McpV<sup>PR</sup>, and enabled ranking of the compounds by affinity. Furthermore, the data support the homology model based on the Adeg\_3718 structure and the involvement of residue Tyr143 in ligand coordination.

### **Glycolate is present in alfalfa seed exudates**

The discovery that carboxylates are sensed by McpV led us to question if they are exuded by germinating alfalfa seeds. We first used a global metabolite profiling platform (UPLC-QTOF) to determine if the carboxylates acetate, propionate, pyruvate, butyrate, glycolate, acetoacetate, and/or isobutyrate were present in the seed exudate. This analysis identified glycolate as the only carboxylate of interest detectable in the seed exudate, and we proceeded to quantify the amount of glycolate in the seed exudate. Our resulting UPLC-MS analysis showed that alfalfa seed exudates contain  $290 \pm 94$  pmol/seed. With an average seed volume of 2.17  $\mu\text{l}$ , the concentration of

glycolate at the surface of the seed is calculated to be  $132 \pm 42 \mu\text{M}$  (15). This concentration of glycolate on the seed surface is relevant for chemotaxis (Fig. 3.3A).

## DISCUSSION

Chemotaxis has been thoroughly established as a critical facet of nodule occupancy and competition in symbiotic rhizobacteria (15-25, 35). Plants exude a plethora of compounds such as amino acids, sugars, organic acids, flavonoids, lipids, and ions (20-21, 36-38). The sensory repertoire of a bacterium as mediated by MCPs encompasses the range of compounds that are important to its lifestyle. Logically, the sensing profiles of root-associated organisms should evolve around the exudation profiles of their respective hosts.

The traditional capillary assay is a robust method of quantifying and comparing chemotaxis responses to different attractants (34). Using this technique, we identified and compared seven new attractants of *S. meliloti* (Fig. 3.3). It should be noted that because of diffusion, the bacteria in the pond are sensing a concentration that is always less than what is loaded into the capillary (39). Acetate and propionate recruited the largest number of cells caused the greatest shift in  $T_m$  of  $\text{McpV}^{\text{PR}}$  in the DSF assay and had the tightest interactions each as determined by ITC with a  $K_d$  in the micromolar range (Figs. 3.2, 3.7B, 3.7C, Table 3.1). Glycolate and pyruvate drew slightly fewer cells to the capillary. The  $T_m$  of  $\text{McpV}^{\text{PR}}$  in the presence of pyruvate was closer to that of acetate and propionate, while glycolate was slightly lower than the previous three (Fig. 3.2, Table 3.1). The  $K_d$  values of glycolate and pyruvate were similar, in the ten micromolar range (Figs. 3.7E, 3.7F). Acetoacetate recruited bacteria in quantities similar to acetate and propionate. The concentration eliciting peak chemotaxis, however, was 100 times greater than for acetate and propionate (100 mM versus 1 mM) (Fig. 3.3). The  $\Delta T_m$  of  $\text{McpV}^{\text{PR}}$  in the presence of acetoacetate was two-fold lower than those in the presence of acetate and propionate. Correspondingly, the  $K_d$  of acetoacetate was the second highest among the six that were determined, in the hundred micromolar range (Figs. 3.2A, 3.7G). Formate drew relatively few cells at its peak concentration of attraction, which is likely a result of its comparatively weak  $K_d$  of about 8.7 mM (Figs. 3.2A, 3.7A) It should be noted that ITC is unfit to monitor interactions with affinities higher than 10 mM (40). Butyrate also elicited one of the lowest attractant responses (Fig. 3.2B). The affinity of this molecule and its isomer, isobutyrate, to  $\text{McpV}$  could not be determined because of apparent

conflicting interactions detected in the ITC experiments. It is possible that the protein construct interacts with those molecules by a mechanism that is non-physiologically relevant, in addition to an association in the canonical binding pocket (Figs. 3.7D, 3.7H). Interestingly, a titration with 10 mM trichloroacetate also produced a similar multiphasic isotherm (data not shown). When comparing the behavioral and *in-vitro* binding data, attractant strength does not necessarily correlate with  $K_d$ . Instead, it appears that  $K_d$  matches more closely with peak concentration of attraction. Similarly, when the PR of a *Pseudomonas aeruginosa* chemoreceptor was fused to the signaling domain of the Tar chemoreceptor to create a chimera in *E. coli*, correlations were found between ligand affinity and signal output or attractant utilization using a FRET assay (41). However, this pattern does not hold for all systems, such as the *Pseudomonas putida* chemoreceptors for TCA intermediates (McpS) and cyclic carboxylates (PcaY\_PP) that do not show a distinct difference in ligand-binding affinity for differently utilized attractants (42, 43). In conclusion, small carboxylates are a new class of attractants for *S. meliloti* that are directly sensed by McpV.

Combining DSF with Biolog PM plates is an effective and facile method for screening the ligand profile of a chemoreceptor (30, 33). While not as robust as ITC, this technique has the advantage of being high throughput and does not require large amounts of protein or ligand. A significant  $\Delta T_m$  was determined to be 3°C because it was clearly larger than the  $\Delta T_m$  of McpV<sup>PR</sup> in the presence of most other compounds (Fig. 3.2). This boundary does not appear to define whether or not a compound serves as a ligand. When screening a single-point variant of McpV<sup>PR</sup>, significant  $T_m$  shifts were still identified, while ITC data showed that the variant protein bound to propionate with a 1,000-fold lower affinity, but did not interact with isobutyrate (Figs. 3.7C, 3.8, Table 3.1; data not shown). Studies of protein-ligand interactions *in vitro* suggest that the minimum requirement for binding to McpV is a carboxylate group. In summary, DSF is an excellent first screen for the putative ligands of proteins, but requires validation through other *in-vitro* studies. Formate, butyrate, and isobutyrate had weak or unorthodox interactions with McpV<sup>PR</sup>, which is supported by data showing that chemotaxis to the former two required a copy of *mcpV* (Figs. 3.5A, 3.7A, 3.7D, 3.7H). Both four-carbon carboxylates are likely too large to properly fit into the binding pocket of McpV. Formate interacts weakly with McpV, while butyrate and isobutyrate may interact with McpV in a more complicated manner. Together, these data indicate that formate and butyrate are very ineffective attractants.

Homologues of McpV have been characterized in two separate species of *Pseudomonas* (29, 30). McpP of *Pseudomonas putida* KT2440 was reported to mediate taxis to and directly bind L-lactate, acetate, propionate, and pyruvate. Propionate and pyruvate elicited the highest magnitude of chemotaxis. In ITC studies propionate, pyruvate and acetate all bound to McpP with very similar  $K_d$  values between 30 and 40  $\mu\text{M}$  (29). In *Pseudomonas syringae* pv. *actinidae*, PscD was characterized as a small carboxylate chemoreceptor using capillary assays, ITC, and protein crystallography. The dissociation constants for glycolate, acetate, propionate, and pyruvate were 23, 31, 101, and 356  $\mu\text{M}$ , respectively. Glycolate had the highest attractant response in this study, while the other three attractants examined were found to draw similar numbers of bacteria. Neither study tested acetoacetate chemotaxis or binding of the *Pseudomonas* chemoreceptors to acetoacetate. Therefore, the homologue in *S. meliloti* is the only known acetoacetate chemoreceptor. The sensor domain of PscD was crystallized in the presence of propionate, defining the ligand binding pocket and coordination sites (30). Both *Pseudomonas* sensors contain the conserved residues involved in ligand binding as identified in Fig. 1B for McpV. The comparison of these three homologues begs the question of why homologous sensors in the respective organisms have different preferences. Another interesting avenue to investigate is the structural basis for the differences in chemotactic potency of an attractant and *in-vitro* ligand affinity.

Navigation to the root is a key first step in the interaction between *S. meliloti* and its host, alfalfa (44, 45). Only glycolate, but none of the other carboxylate attractants tested, was detected in the exudate of alfalfa seeds. We predict that pyruvate and acetoacetate are either not exuded or were not detected because of their instability. UPLC-MS revealed that the concentration of glycolate at the surface of a germinating alfalfa seed is within the sensing range of *S. meliloti* (Fig. 3.3A). We previously reported that alfalfa seedlings also exude proline and choline in sufficient quantities to be effectively sensed by McpU and McpX, the respective sensors in *S. meliloti* (16, 17). Taking into account the exudation profile of alfalfa seedlings, it appears that short chain carboxylates are not a major avenue of host seed sensing. The exudation of the small carboxylates sensed by McpV in different spatiotemporal contexts should not be ruled out. Acetate, formate, and lactate were detected in the root exudates of two species of the legume genus *Lupinus* during both flowering and fruiting periods (46).



McpV is clearly a critical chemoreceptor because we recently determined that it is the most abundant chemoreceptor in *S. meliloti*, accounting for 70% of the total pool of chemoreceptors (47). Perhaps this explains the 1.5-fold increase in chemotaxis to proline in the  $\Delta mcpV$  strain. The elimination of McpV from the chemoreceptor array could result in an overrepresentation of the remaining chemoreceptors and their respective signals. Most of the McpV ligands analyzed are not exuded by alfalfa seeds, so their purpose in *S. meliloti* may not be limited to host-microbe interaction. We hypothesize that taxis to these carbon sources is critical to the survival of the bacterium in the bulk soil, which contains many different organic acids (48-50). Acetate, acetoacetate, and propionate have been identified as utilizable carbon sources by *S. meliloti* (51-53). Interestingly, formate acts as an electron donor during the chemoautotrophic growth of *S. meliloti* on carbonate (54). The genomes of *S. meliloti* 1021 and RU11/001 have a putative *glcDEF* operon, which may allow the use of glycolate in the glyoxylate shunt (55-58).

The characterization of McpV adds a new class of compounds to the known sensory repertoire of *S. meliloti*. Currently, this includes proteogenic and non-proteogenic amino acids, quaternary ammonium compounds, and two to four carbon carboxylates. The function of the remaining five receptors remains to be elucidated. McpY exhibits similarity to receptors involved in energy taxis – the phenomenon where bacteria accumulate in regions rich in compounds that can act as electron donors, and IcpA is annotated to contain a HemAT domain, which is involved in sensing oxygen (5, 23, 59, 60). The periplasmic regions of McpT, McpW, and McpZ remain to be annotated and the receptors have yet to be characterized. The range of compounds sensed by an individual chemoreceptor can be expanded through indirect sensing via interaction with periplasmic binding proteins. In *E. coli*, maltose-bound maltose binding protein interacts with Tar, which permits maltose taxis. *Bacillus subtilis* senses multiple amino acids via indirect and direct binding to McpC (61-63). The elucidation of chemoreceptor function in attractant sensing will increase our knowledge of plant-microbe interactions and bacterial lifestyles. Root exudates are a major avenue for plants and microbes to exchange nutrients and information. Understanding the establishment and maintenance of microbial communities in the rhizosphere is a critical objective of improving modern agriculture.

## MATERIALS AND METHODS

### **Strains and plasmids**

*S. meliloti* strains are highly motile MVII-1 derivatives and listed in Table 2. Derivatives of *E. coli* K12 strains and plasmids used for molecular techniques are also listed in Table 2.

### **Media and growth conditions**

*E. coli* was grown using lysogeny broth (LB) at 37°C (64). TYC medium was used to grow *S. meliloti* at 30°C and contained 0.5% tryptone, 0.3% yeast extract (BD, Sparks, MD), and 6 mM CaCl<sub>2</sub> (Fischer, Fairlawn, NJ) with 600 µg/ml streptomycin. Minimal medium used for *S. meliloti* was Rhizobium Basal medium (RB) and contained 0.1 mM NaCl, 0.01 Na<sub>2</sub>MoO<sub>4</sub>, 6.1 mM K<sub>2</sub>HPO<sub>4</sub>, 3.9 mM KH<sub>2</sub>PO<sub>4</sub>, 1 mM (NH<sub>4</sub>)<sub>2</sub>SO<sub>4</sub>, 1 µM FeSO<sub>4</sub>, 1 mM MgSO<sub>4</sub>, 0.1 mM CaCl<sub>2</sub>, 20 µg/l D-biotin, and 10 µg/l thiamin (65). Low nutrient Bromfield plates were prepared according to Sourjik and Schmitt (66). Ampicillin and kanamycin concentrations used were 100 µg/ml and 25 µg/ml, respectively. Authentic organic acid standards were purchased from Supelco (Bellefonte, PA), except for lithium acetoacetate and glycolic acid, which were supplied from TCI (Tokyo, Japan).

### **Preparation of seed exudates**

An amount of 0.1 g *M. sativa* Guardsman II variety seeds were rinsed four times with sterile water and then soaked in 3% H<sub>2</sub>O<sub>2</sub> for 12 min. The seeds were rinsed four more times with sterile water and placed into a 125 ml Erlenmeyer flask with 3 ml of sterile water. Seeds were examined by eye, and exudates were viewed under a microscope for contamination. An amount of 200 µl exudate was plated onto TYC to check for contamination. Samples that appeared visually clear were flash frozen in liquid nitrogen and stored at -80°C. If no growth was observed on the TYC plates the following day, samples were thawed on ice, sonicated for 10 min in 30-40 s pulses, centrifuged at 5,000 × g for 10 min, and supernatants were withdrawn to yield seed exudates.

### **Quantification of organic acids in seed exudate**

Global metabolite profiling to determine if carboxylates were present in the seed exudate was performed on a Waters I-class Acquity UPLC interfaced with a Waters Synapt G2-S mass spectrometer operated in high resolution mode (Waters Corp., Milford, MA). The UPLC was fitted with a Waters BEH c18 column (1.7 µm, 2.1x 50 mm) and the mobile phase A consisted of water + 0.1% formic acid and mobile phase B consisted of acetonitrile + 0.1% formic acid. The flow rate

was 0.2 ml/min and the 10 minute gradient was initial – 2 minutes 0.5%B, 6 minutes 10% B, 8.5 minutes 90%B, 9-10 minutes equilibration at 0.5%B. Mass spectrometer data collection was performed in both positive and negative mode with a mass range of 50-1800, capillary voltage of 2.2 kV, cone voltage at 10, desolvation gas flow 450 L/Hr and cone gas flow 45 L/Hr. Presence and absence of the carboxylates was determined by extracting the ion mass of the carboxylate from the total ion chromatogram ([M-H]<sup>-</sup> glycolate 75.0088, pyruvate 87.0088, butyrate and isobutyrate 87.0452, acetate 61.0284, propionate 73.0295 and [M+H]<sup>+</sup> acetoacetate 103.0390).

Glycolic acid quantification was performed using a Waters H-class Acquity UPLC interfaced with Waters Xevo-MS mass spectrometer (Waters Corp., Milford, MA). The UPLC was equipped with a Rezex ROA-Organic Acid column maintained at 55°C with a mobile phase of water with 0.5% formic acid at a flow rate of 0.25 ml/min for 10 min. The mass spectrometer was operated in SIR mode with unit resolution set to detect at 75.0 m/z, cone voltage 24 and dwell time 1.15 seconds. A glycolic acid standard was purchased from Supelco (Bellefonte, PA) and used to establish a calibration curve from 0.25 – 2.5 µg/ml.

### **Mutant construction and genetic manipulation**

Single point mutations in *mcpV* were made *in vitro* using overlap-extension PCR (67). Allelic exchange mutagenesis was used to construct markerless mutants according to previous protocols (68, 69). DNA isolation and cleanup were performed with Wizard kits from Promega according to manufacturer's instructions.

### **Capillary assay**

Capillary assays were performed as originally described by Adler (34) with minor modifications for *S. meliloti* (17). Motile *S. meliloti* cells were obtained by diluting stationary phase TYC cultures into 10 ml of RB overlain onto Bromfield agar plates and incubating at 30°C for 15 h. Cells were harvested between an OD<sub>600</sub> of 0.16 and 0.18 and sedimented by centrifugation at 3,000 × g for 5 min before being suspended to a final OD<sub>600</sub> of 0.15. A culture amount of 375 µl was placed into a pond formed from a U-shaped glass tube between two glass plates. One-µl Microcaps glass capillaries (Drummond Scientific, Broomall, PA) were sealed at one end over a flame and placed into a ligand solution in a vacuum chamber. A vacuum was created in the chamber to allow the

solution to fill the capillary after the air was removed and the vacuum was released. Capillaries were placed into the bacterial ponds and left to incubate at room temperature for 2 h. The capillaries were then removed, broken at the sealed tip, and their contents expelled into RB. Serial dilutions were plated in duplicates onto TYC containing streptomycin plates, and colonies were counted after three days of growth. The counts of a control capillary were subtracted from all test capillaries to account for accumulation due to random movement of bacteria into the capillary. Three technical replicates were performed for each of three biological replicates.

### **Homology modeling**

To construct the model of McpV-PR, the amino acid sequence of McpV between Gln33 and Gln189 was uploaded to the SWISS-MODEL server (Swiss Institute of Bioinformatics) (70-76). The template used was PDB ID 4k08, a crystallized product of recombinant Adeh\_3718, from the soil bacterium *Anaeromyxobacter dehalogenans* (32).

### **Over-expression and purification of McpV<sup>PR</sup>**

*E. coli* M15/pREP4 was transformed with pBS377, and expression cultures were grown to an OD<sub>600</sub> between 0.7 and 0.9 before induction with 0.6 mM isopropyl-thiogalactopyranoside. Cultures were further incubated either for 4 h at 25°C, or for 16 h at 16°C. Cells were harvested by centrifugation at 9,500 × g for 9 min at 4°C. Cell pellets were suspended in a binding buffer consisting of 0.5 M NaCl, 20 mM imidazole, 1 mM phenylmethane sulfonyl fluoride and 20 mM sodium phosphate, pH 7.4. The cells were lysed by two to three passages through a French pressure cell at 16,000 lb/in<sup>2</sup> (SLM Aminco, Silver Spring, MD). Lysates were centrifuged at 56,000 × g for 50 min at 4°C, followed by filter sterilization. The clarified lysate was then applied to a nickel-NTA affinity column (GE Healthcare), and the column was washed with binding buffer. To elute the protein, an elution buffer composed of 0.5 M NaCl, 20 mM sodium phosphate, 0.5 M imidazole, and a pH of 7.4 was applied to the column in an increasing linear gradient. Protein elution was monitored by UV absorbance and confirmed by SDS-PAGE. Pooled fractions containing protein were further purified by size exclusion chromatography using a HiPrep 26/60 Sephacryl S-300 HR column (GE Healthcare) in 100 mM NaCl, 25 mM HEPES, pH 7.0 for DSF experiments, or 0.4 M NaCl, 25 mM HEPES, pH 8.0 for ITC experiments. When appropriate, the protein was concentrated using an Amicon ultrafiltration system and regenerated cellulose

membranes (Millipore, Billerica, MA). Protein concentration was determined using UV spectrometry and a theoretical extinction coefficient of  $37,930 \text{ M}^{-1} \text{ cm}^{-1}$  obtained from the ExPASy online ProtParam tool (77).

### **Differential scanning fluorimetry (DSF)**

Compounds in Biolog MicroPlate PM1 (Biolog, Hayward, CA) were dissolved in a master mix of  $10 \mu\text{M}$  McpV<sup>PR</sup> and 1.4x SYPRO Orange in the same size exclusion buffer used to purify the protein. According to the manufacturer each well contains 0.5 to 1  $\mu\text{moles}$  of compound, making the final concentrations between 7 and 15 mM. A volume of 30  $\mu\text{l}$  from each well was transferred to a 96-well plate reader for use in an ABI 7300 real PCR system. The temperature gradient began at  $10^\circ\text{C}$  and increased in  $0.5^\circ\text{C}$  steps every 30 s to  $90^\circ\text{C}$ . The melting temperature ( $T_m$ ) of the protein in each well was defined as the peak of the first derivative of the fluorescence curve. The melting temperature shift ( $\Delta T_m$ ) was determined by subtracting the  $T_m$  of the control well containing no ligand from the  $T_m$  of each test well. The screen was performed in triplicate using three Biolog plates. For the mutant protein, the screen was performed once.

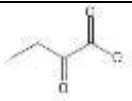
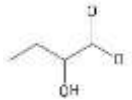
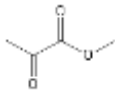
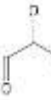
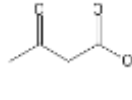
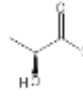
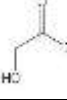
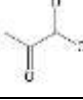
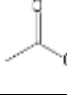
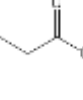
### **Isothermal titration calorimetry (ITC)**

Direct binding studies were performed with a VP-ITC microcalorimeter MicroCal (Malvern, Westborough, MA). McpV<sup>PR</sup> was used at  $75 \mu\text{M}$  and titrated against 2 to 15 mM of each carboxylate. The experiment was performed at  $25^\circ\text{C}$  for all compounds except acetoacetate, which was performed at  $28^\circ\text{C}$ . Prior to experiments, both protein and ligand solution were degassed at a temperature 2 to  $3^\circ\text{C}$  above the experimental temperature. All ligand solutions were made with the same batch of 0.4 M NaCl, 25 mM HEPES, pH 8.0 used for the final protein purification step. For baseline titrations, the ligand was titrated into the buffer without protein, which was used a reference subtraction for the respective titrations with protein. Association constants were reported from curves generated using the MicroCal version of Origin 7.0 software using the “one binding sites” model. (Origin Lab, Northampton, MA).

### **ACKNOWLEDGEMENTS**

This study was supported by NSF grant MCB-1253234 to Birgit Scharf. We are indebted to Florian Schubot for sharing instrument ABI 7300 real-time PCR system and support with protein modeling. The Virginia Tech Mass Spectrometry Incubator is maintained with funding from the Fralin Life Science Institute of Virginia Tech as well as NIFA (Hatch Grant 228344 and VA-160085).

**Table 3.1.** The ten compounds that caused the greatest shift in melting temperature of McpV<sup>PR</sup> in the thermal denaturation assay. The structure and length of the carbon chain is provided for comparison. Compounds are sorted according to ascending T<sub>m</sub> values of the wild-type protein.

Compound	McpV <sup>PR</sup> ΔT <sub>m</sub>	McpV <sup>Y143A-PR</sup> ΔT <sub>m</sub>	Structure	Number of carbons
<b>α-keto-butyrate</b>	3.8	4.5		4
<b>α-hydroxy-butyrate</b>	4	1.5		4
<b>methyl-pyruvate</b>	5.8	4		3
<b>glyoxylate</b>	5.8	4.5		2
<b>acetoacetate</b>	6	6		4
<b>L-lactate</b>	8.2	2		3
<b>glycolate</b>	9.3	6		2
<b>pyruvate</b>	11.8	9.5		3
<b>acetate</b>	12.3	9		2
<b>propionate</b>	12.3	9		3

**Table 3.2.** Bacterial strains and plasmids

<b>Strain</b>	<b>Characteristics</b>	<b>Reference/Source</b>
<i>E. coli</i>		
DH5 $\alpha$	<i>recA1 endA1</i>	(78)
M15/pREP4	Km <sup>r</sup> , expression strain <i>lac mtl</i>	Qiagen
S17-1	<i>recA endA thi hsdR</i> RP4-2 Tc::Mu::Tn7 Tp <sup>r</sup> Sm <sup>r</sup>	(68)
<i>S. meliloti</i>		
RU11/001	Sm <sup>r</sup> , spontaneously streptomycin-resistant wild-type strain	(79)
RU11/830	Sm <sup>r</sup> , $\Delta mcpV$	(23)
RU13/149	Sm <sup>r</sup> , $\Delta mcpS$ , $\Delta mcpT$ , $\Delta mcpU$ , $\Delta mcpV$ , $\Delta mcpW$ , $\Delta mcpX$ , $\Delta mcpY$ , $\Delta mcpZ$ , $\Delta icpA$	(23)
BS232	Sm <sup>r</sup> , <i>mcpV</i> -Y143A	This work
BS234	Sm <sup>r</sup> , <i>mcpV</i> -H103D	This work
Plasmids		
pK18 <i>mobsacB</i>	Km <sup>r</sup> , <i>lacZ mob sacB</i>	(80)
pQE60	Ap <sup>r</sup> , expression vector	Qiagen
pBS377	Ap <sup>r</sup> , pQE60 with <i>mcpV</i> 33 to 471 bp <i>NcoI/BamHI</i> PCR fragment containing <i>mcpV</i> 96-567 bp (aa 33-189)	This work
pBS1151	Ap <sup>r</sup> , pQE60 with <i>mcpV</i> 33 to 471 bp <i>NcoI/BamHI</i> PCR fragment containing <i>mcpV</i> 96-567 bp (aa 33-189) with Y143A substitution	This work



## REFERENCES

1. Hazelbauer GL, Falke JJ, Parkinson JS. 2008. Bacterial chemoreceptors: high-performance signaling in networked arrays. *Trends Biochem Sci* 33:9-19.
2. Parkinson JS, Hazelbauer GL, Falke JJ. 2015. Signaling and sensory adaptation in *Escherichia coli* chemoreceptors: 2015 update. *Trends Microbiol* 23:257-66.
3. Bi S, Lai L. 2015. Bacterial chemoreceptors and chemoeffectors. *Cell Mol Life Sci* 72:691-708.
4. Krell T, Lacal J, Munoz-Martinez F, Reyes-Darias JA, Cadirci BH, Garcia-Fontana C, Ramos JL. 2011. Diversity at its best: bacterial taxis. *Environ Microbiol* 13:1115-24.
5. Miller LD, Russell MH, Alexandre G. 2009. Diversity in bacterial chemotactic responses and niche adaptation. *Adv Appl Microbiol* 66:53-75.
6. Lacal J, Garcia-Fontana C, Munoz-Martinez F, Ramos JL, Krell T. 2010. Sensing of environmental signals: classification of chemoreceptors according to the size of their ligand binding regions. *Environ Microbiol* 12:2873-84.
7. Scharf BE, Hynes MF, Alexandre GM. 2016. Chemotaxis signaling systems in model beneficial plant-bacteria associations. *Plant Molecular Biology* 90:549-559.
8. Brewin NJ. 1991. Development of the legume root nodule. *Annu Rev Cell Biol* 7:191-226.
9. Kondorosi E, Mergaert P, Kereszt A. 2013. A paradigm for endosymbiotic life: cell differentiation of *Rhizobium* bacteria provoked by host plant factors. *Annu Rev Microbiol* 67:611-28.
10. Haag AF, Arnold MF, Myka KK, Kerscher B, Dall'Angelo S, Zanda M, Mergaert P, Ferguson GP. 2013. Molecular insights into bacteroid development during *Rhizobium*-legume symbiosis. *FEMS Microbiol Rev* 37:364-83.
11. Drogue B, Dore H, Borland S, Wisniewski-Dye F, Prigent-Combaret C. 2012. Which specificity in cooperation between phytostimulating rhizobacteria and plants? *Res Microbiol* 163:500-10.
12. van Rhijn P, Vanderleyden J. 1995. The *Rhizobium*-plant symbiosis. *Microbiol Rev* 59:124-42.
13. Services NAS. 2017. Crop Production 2016 Summary. United States Department of Agriculture,

14. Lindeman WC, Glover CR, Flynn R, Idowu J. June, 2015 2015. Nitrogen Fixation by Legumes, on New Mexico State University Cooperative Extension. [http://aces.nmsu.edu/pubs/\\_a/A129/](http://aces.nmsu.edu/pubs/_a/A129/). Accessed 8/29/2017.
15. Webb BA, Helm RF, Scharf BE. 2016. Contribution of individual chemoreceptors to *Sinorhizobium meliloti* chemotaxis towards amino acids of host and nonhost seed exudates. *Mol Plant Microbe Interact* 29:231-239.
16. Webb BA, Hildreth S, Helm RF, Scharf BE. 2014. *Sinorhizobium meliloti* chemoreceptor McpU mediates chemotaxis toward host plant exudates through direct proline sensing. *Appl Environ Microbiol* 80:3404-15.
17. Webb BA, Compton KK, Castaneda Saldana R, Arapov T, Ray WK, Helm RF, Scharf BE. 2017. *Sinorhizobium meliloti* chemotaxis to quaternary ammonium compounds is mediated by the chemoreceptor McpX. *Mol Microbiol* 103:333-346.
18. Bernabeu-Roda L, Calatrava-Morales N, Cuellar V, Soto MJ. 2015. Characterization of surface motility in *Sinorhizobium meliloti*: regulation and role in symbiosis. *Symbiosis* 67:79-90.
19. Miller LD, Yost CK, Hynes MF, Alexandre G. 2007. The major chemotaxis gene cluster of *Rhizobium leguminosarum* bv. *viciae* is essential for competitive nodulation. *Mol Microbiol* 63:348-62.
20. Moe LA. 2013. Amino acids in the rhizosphere: from plants to microbes. *Am J Bot* 100:1692-705.
21. Barbour WM, Hattermann DR, Stacey G. 1991. Chemotaxis of *Bradyrhizobium japonicum* to soybean exudates. *Appl Environ Microbiol* 57:2635-9.
22. Nelson EB. 2004. Microbial dynamics and interactions in the spermosphere. *Annu Rev Phytopathol* 42:271-309.
23. Meier VM, Muschler P, Scharf BE. 2007. Functional analysis of nine putative chemoreceptor proteins in *Sinorhizobium meliloti*. *J Bacteriol* 189:1816-1826.
24. Meier VM, Scharf BE. 2009. Cellular localization of predicted transmembrane and soluble chemoreceptors in *Sinorhizobium meliloti*. *J Bacteriol* 191:5724-33.
25. Webb BA, Compton KK, Del Campo JSM, Taylor D, Sobrado P, Scharf BE. 2017. *Sinorhizobium meliloti* chemotaxis to multiple amino acids is mediated by the chemoreceptor McpU. *Mol Plant Microbe Interact* 30:770-777.

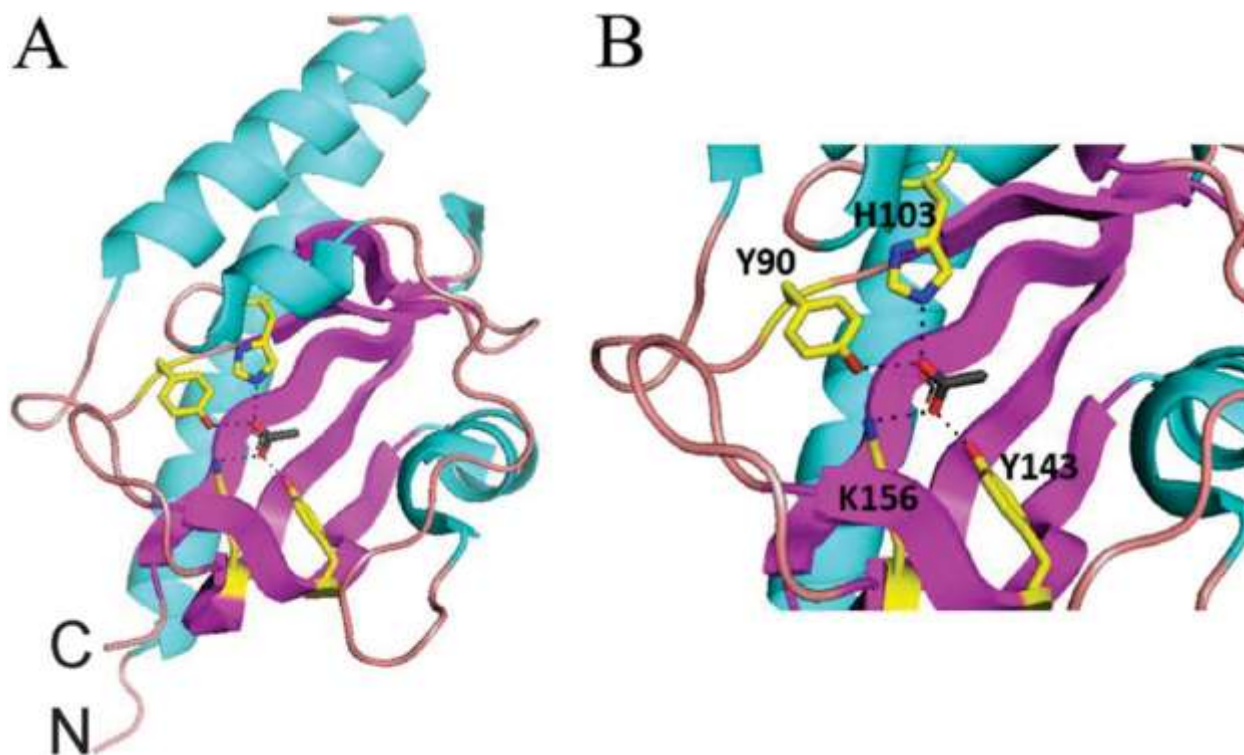
26. Ortega A, Zhulin IB, Krell T. 2017. Sensory repertoire of bacterial chemoreceptors. *Microbiol Mol Biol Rev* 81.
27. Anantharaman V, Aravind L. 2000. Cache - a signaling domain common to animal Ca(2<sup>+</sup>)-channel subunits and a class of prokaryotic chemotaxis receptors. *Trends Biochem Sci* 25:535-537.
28. Upadhyay AA, Fleetwood AD, Adebali O, Finn RD, Zhulin IB. 2016. Cache domains that are homologous to, but different from PAS domains comprise the largest superfamily of extracellular sensors in prokaryotes. *PLoS Comput Biol* 12:e1004862.
29. Garcia V, Reyes-Darias JA, Martin-Mora D, Morel B, Matilla MA, Krell T. 2015. Identification of a chemoreceptor for C2 and C3 carboxylic acids. *Appl Environ Microbiol* 81:5449-57.
30. Brewster JL, McKellar JL, Finn TJ, Newman J, Peat TS, Gerth ML. 2016. Structural basis for ligand recognition by a Cache chemosensory domain that mediates carboxylate sensing in *Pseudomonas syringae*. *Sci Rep* 6:35198.
31. Finn RD, Coggill P, Eberhardt RY, Eddy SR, Mistry J, Mitchell AL, Potter SC, Punta M, Qureshi M, Sangrador-Vegas A, Salazar GA, Tate J, Bateman A. 2016. The Pfam protein families database: towards a more sustainable future. *Nucleic Acids Research* 44:D279-D285.
32. Pokkuluri PR, Dwulit-Smith J, Duke NE, Wilton R, Mack JC, Bearden J, Rakowski E, Babnigg G, Szurmant H, Joachimiak A, Schiffer M. 2013. Analysis of periplasmic sensor domains from *Anaeromyxobacter dehalogenans* 2CP-C: Structure of one sensor domain from a histidine kinase and another from a chemotaxis protein. *Microbiologyopen* 2:766-777.
33. McKellar JLO, Minnell JJ, Gerth ML. 2015. A high-throughput screen for ligand binding reveals the specificities of three amino acid chemoreceptors from *Pseudomonas syringae* pv. *actinidiae*. *Molecular Microbiology* 96:694-707.
34. Adler J. 1973. A method for measuring chemotaxis and use of the method to determine optimum conditions for chemotaxis by *Escherichia coli*. *J Gen Microbiol* 74:77-91.
35. Owen AG, Jones DL. 2001. Competition for amino acids between wheat roots and rhizosphere microorganisms and the role of amino acids in plant N acquisition. *Soil Biology & Biochemistry* 33:651-657.

36. Jones DL, Darrah PR. 1994. Amino-Acid Influx at the Soil-Root Interface of Zea-Mays L and Its Implications in the Rhizosphere. *Plant and Soil* 163:1-12.
37. Jones DL, Edwards AC, Donachie K, Darrah PR. 1994. Role of Proteinaceous Amino-Acids Released in Root Exudates in Nutrient Acquisition from the Rhizosphere. *Plant and Soil* 158:183-192.
38. Odunfa VSA. 1979. Free Amino-Acids in the Seed and Root Exudates in Relation to the Nitrogen Requirements of Rhizosphere Soil *Fusaria*. *Plant and Soil* 52:491-499.
39. Futrelle RP, Berg HC. 1972. Specification of gradients used for studies of chemotaxis. *Nature* 239:517-8.
40. Wiseman T, Williston S, Brandts JF, Lin LN. 1989. Rapid measurement of binding constants and heats of binding using a new titration calorimeter. *Analytical Biochemistry* 179:131-137.
41. Reyes-Darias JA, Yang YL, Sourjik V, Krell T. 2015. Correlation between signal input and output in PctA and PctB amino acid chemoreceptor of *Pseudomonas aeruginosa*. *Molecular Microbiology* 96:513-525.
42. Lacal J, Alfonso C, Liu XX, Parales RE, Morel B, Conejero-Lara F, Rivas G, Duque E, Ramos JL, Krell T. 2010. Identification of a Chemoreceptor for Tricarboxylic Acid Cycle Intermediates. *Journal of Biological Chemistry* 285:23124-23134.
43. Fernandez M, Matilla MA, Ortega A, Krell T. 2017. Metabolic Value Chemoattractants Are Preferentially Recognized at Broad Ligand Range Chemoreceptor of *Pseudomonas putida* KT2440. *Frontiers in Microbiology* 8.
44. Ames P, Bergman K. 1981. Competitive Advantage Provided by Bacterial Motility in the Formation of Nodules by *Rhizobium-meliloti*. *Journal of Bacteriology* 148:728-729.
45. Gulash M, Ames P, Larosiliere RC, Bergman K. 1984. Rhizobia Are Attracted to Localized Sites on Legume Roots. *Applied and Environmental Microbiology* 48:149-152.
46. Garcia JAL, Barbas C, Probanza A, Barrientos ML, Manero FJG. 2001. Low molecular weight organic acids and fatty acids in root exudates of two *Lupinus* cultivars at flowering and fruiting stages. *Phytochemical Analysis* 12:305-311.
47. Zatakia HM, Arapov TD, Meier VM, Scharf BE. 2018. Cellular Stoichiometry of Methyl-Accepting Chemotaxis Proteins in *Sinorhizobium meliloti*. *Journal of Bacteriology* 200.

48. Cieslinski G, Van Rees KCJ, Szmigielska AM, Krishnamurti GSR, Huang PM. 1998. Low-molecular-weight organic acids in rhizosphere soils of durum wheat and their effect on cadmium bioaccumulation. *Plant and Soil* 203:109-117.
49. van Hees PAW, Dahlen J, Lundstrom US, Boren H, Allard B. 1999. Determination of low molecular weight organic acids in soil solution by HPLC. *Talanta* 48:173-179.
50. Li XL, Chen XM, Liu X, Zhou LC, Yang XQ. 2012. Characterization of soil low-molecular-weight organic acids in the Karst rocky desertification region of Guizhou Province, China. *Frontiers of Environmental Science & Engineering* 6:195-203.
51. Charles TC, Cai GQ, Aneja P. 1997. Megaplasmid and chromosomal loci for the PHB degradation pathway in *Rhizobium (Sinorhizobium) meliloti*. *Genetics* 146:1211-1220.
52. Dunn MF. 1998. Tricarboxylic acid cycle and anaplerotic enzymes in rhizobia. *Fems Microbiology Reviews* 22:105-123.
53. Biondi EG, Tatti E, Comparini D, Giuntini E, Mocali S, Giovannetti L, Bazzicalupo M, Mengoni A, Viti C. 2009. Metabolic Capacity of *Sinorhizobium (Ensifer) meliloti* Strains as Determined by Phenotype MicroArray Analysis. *Applied and Environmental Microbiology* 75:5396-5404.
54. Pickering BS, Oresnik IJ. 2008. Formate-dependent autotrophic growth in *Sinorhizobium meliloti*. *Journal of Bacteriology* 190:6409-6418.
55. Pellicer MT, Badia J, Aguilar J, Baldoma L. 1996. glc locus of *Escherichia coli*: Characterization of genes encoding the subunits of glycolate oxidase and the glc regulator protein. *Journal of Bacteriology* 178:2051-2059.
56. Lord JM. 1972. Glycolate Oxidoreductase in *Escherichia-Coli*. *Biochimica Et Biophysica Acta* 267:227-&.
57. Capela D, Barloy-Hubler F, Gouzy J, Bothe G, Ampe F, Batut J, Boistard P, Becker A, Boutry M, Cadieu E, Dreano S, Gloux S, Godrie T, Goffeau A, Kahn D, Kiss E, Lelaure V, Masuy D, Pohl T, Portetelle D, Puhler A, Purnelle B, Ramsperger U, Renard C, Thebault P, Vandenberg M, Weidner S, Galibert F. 2001. Analysis of the chromosome sequence of the legume symbiont *Sinorhizobium meliloti* strain 1021. *Proceedings of the National Academy of Sciences of the United States of America* 98:9877-9882.

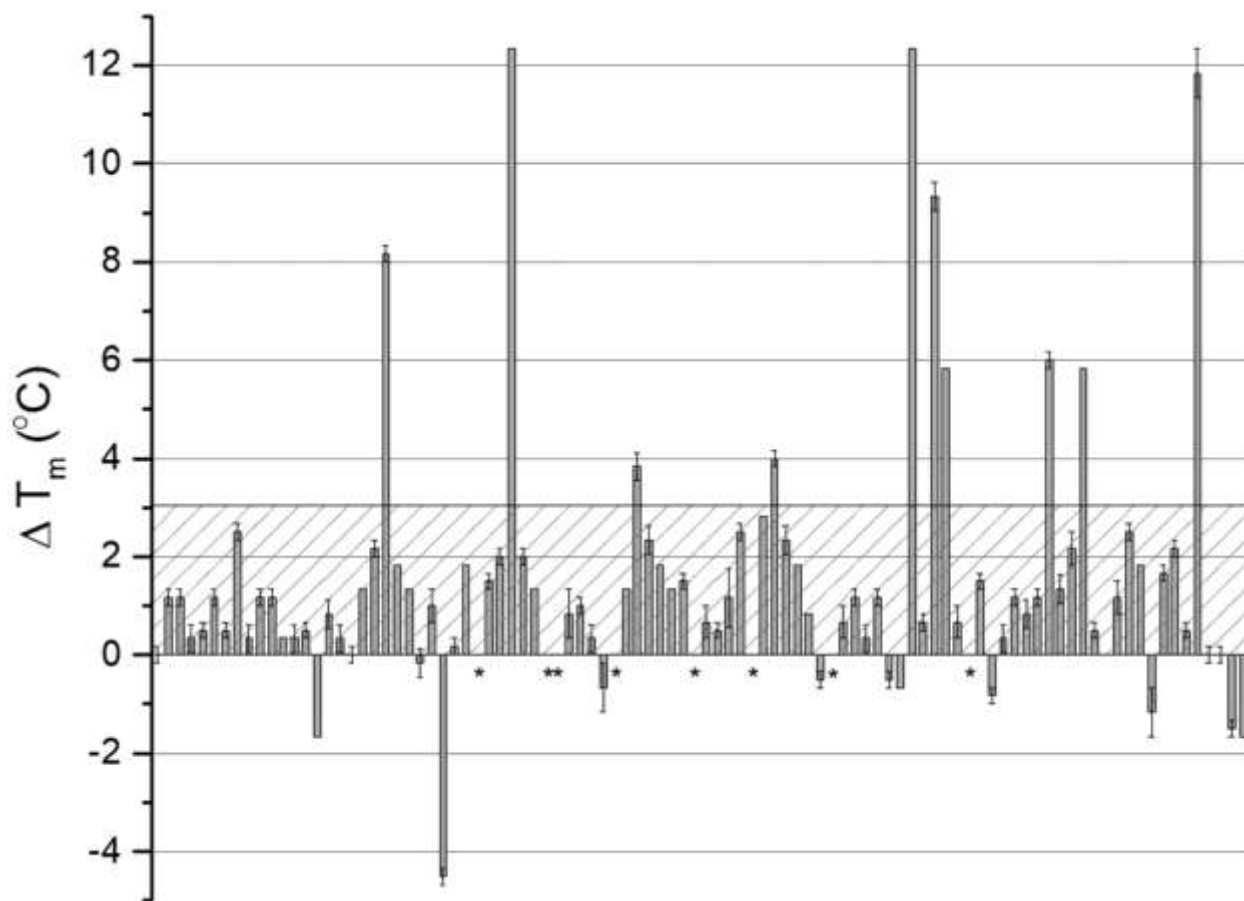
58. Wibberg D, Blom J, Ruckert C, Winkler A, Albersmeier A, Puhler A, Schluter A, Scharf BE. 2013. Draft genome sequence of *Sinorhizobium meliloti* RU11/001, a model organism for flagellum structure, motility and chemotaxis. *J Biotechnol* 168:731-3.
59. Schweinitzer T, Josenhans C. 2010. Bacterial energy taxis: a global strategy? *Archives of Microbiology* 192:507-520.
60. Hou SB, Larsen RW, Boudko D, Riley CW, Karatan E, Zimmer M, Ordal GW, Alam M. 2000. Myoglobin-like aerotaxis transducers in Archaea and Bacteria. *Nature* 403:540-544.
61. Glekas GD, Mulhern BJ, Kroc A, Duelfer KA, Lei V, Rao CV, Ordal GW. 2012. The *Bacillus subtilis* Chemoreceptor McpC Senses Multiple Ligands Using Two Discrete Mechanisms. *Journal of Biological Chemistry* 287.
62. Richarme G. 1982. Interaction of the maltose-binding protein with membrane vesicles of *Escherichia coli*. *J Bacteriol* 149:662-7.
63. Koiwai O, Hayashi H. 1979. Studies on bacterial chemotaxis. IV. Interaction of maltose receptor with a membrane-bound chemosensing component. *J Biochem* 86:27-34.
64. Bertani G. 1951. Studies on Lysogenesis .1. The Mode of Phage Liberation by Lysogenic *Escherichia-Coli*. *Journal of Bacteriology* 62:293-300.
65. Götz R, Limmer N, Ober K, Schmitt R. 1982. Motility and Chemotaxis in 2 Strains of *Rhizobium* with Complex Flagella. *Journal of General Microbiology* 128:789-798.
66. Sourjik V, Schmitt R. 1996. Different roles of CheY1 and CheY2 in the chemotaxis of *Rhizobium meliloti*. *Molecular Microbiology* 22:427-436.
67. Bryksin AV, Matsumura I. 2010. Overlap extension PCR cloning: a simple and reliable way to create recombinant plasmids. *Biotechniques* 48:463-465.
68. Simon R, O'Connell M, Labes M, Pühler A. 1986. Plasmid vectors for the genetic analysis and manipulation of rhizobia and other gram-negative bacteria. *Methods Enzymol* 118:640-659.
69. Simon R, Priefer U, Pühler A. 1983. A broad host range mobilisation system for *in vivo* genetic engineering: Transposon mutagenesis in gram negative bacteria. *Bio/Technology* 1:783-791.
70. Biasini M, Bienert S, Waterhouse A, Arnold K, Studer G, Schmidt T, Kiefer F, Gallo Cassarino T, Bertoni M, Bordoli L, Schwede T. 2014. SWISS-MODEL: modelling protein

- tertiary and quaternary structure using evolutionary information. *Nucleic Acids Res* 42:W252-8.
71. Arnold K, Bordoli L, Kopp J, Schwede T. 2006. The SWISS-MODEL workspace: a web-based environment for protein structure homology modelling. *Bioinformatics* 22:195-201.
  72. Peitsch MC, Schwede T, Guex N. 2000. Automated protein modelling--the proteome in 3D. *Pharmacogenomics* 1:257-66.
  73. Kiefer F, Arnold K, Kunzli M, Bordoli L, Schwede T. 2009. The SWISS-MODEL Repository and associated resources. *Nucleic Acids Res* 37:D387-92.
  74. Biasini M, Bienert S, Waterhouse A, Arnold K, Studer G, Schmidt T, Kiefer F, Cassarino TG, Bertoni M, Bordoli L, Schwede T. 2014. SWISS-MODEL: modelling protein tertiary and quaternary structure using evolutionary information. *Nucleic Acids Research* 42:W252-W258.
  75. Guex N, Peitsch MC, Schwede T. 2009. Automated comparative protein structure modeling with SWISS-MODEL and Swiss-PdbViewer: A historical perspective. *Electrophoresis* 30:S162-S173.
  76. Kiefer F, Arnold K, Kunzli M, Bordoli L, Schwede T. 2009. The SWISS-MODEL Repository and associated resources. *Nucleic Acids Research* 37:D387-D392.
  77. Gasteiger E, Hoogland C, Gattiker A, Duvaud S, Wilkins MR, Appel RD, Bairoch A. 2005. Protein identification and analysis tools on the ExPASy server. In Walker JM (ed), *The proteomics handbook*, Humana Press, Totowa, NJ:571-607.
  78. Hanahan D, Meselson M. 1983. Plasmid screening at high colony density. *Methods Enzymol* 100:333-42.
  79. Pleier E, Schmitt R. 1991. Expression of two *Rhizobium meliloti* flagellin genes and their contribution to the complex filament structure. *J Bacteriol* 173:2077-2085.
  80. Schäfer A, Tauch A, Jager W, Kalinowski J, Thierbach G, Pühler A. 1994. Small mobilizable multi-purpose cloning vectors derived from the *Escherichia coli* plasmids pK18 and pK19: selection of defined deletions in the chromosome of *Corynebacterium glutamicum*. *Gene* 145:69-73.

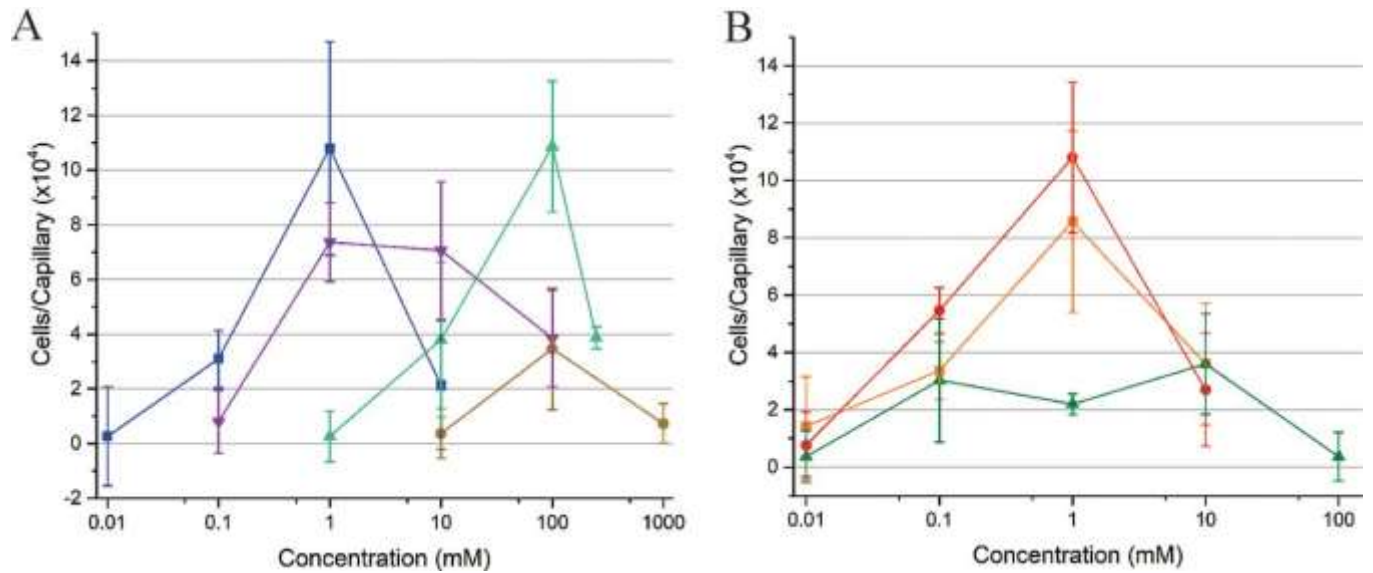


**Fig. 3.1.** Homology model of McpV<sup>PR</sup> using Adeh\_3718 (PDB ID 4k08) as a template. Sequence identity over the modelled range is 54% and the GMQE value is 0.75. **A.** Whole model view. C: C-terminus; N: N-terminus **B.** Close-up view of the binding pocket displaying acetate coordination. Residues in close proximity to the ligand are drawn with yellow carbon chains, red oxygen atoms, and blue nitrogen atoms. The dotted lines indicate possible ligand coordinating bonds to Y90 and H103, which are closest to the upper half of the carboxylate, while K156 and Y143 are closer to the lower half of the carboxylate.

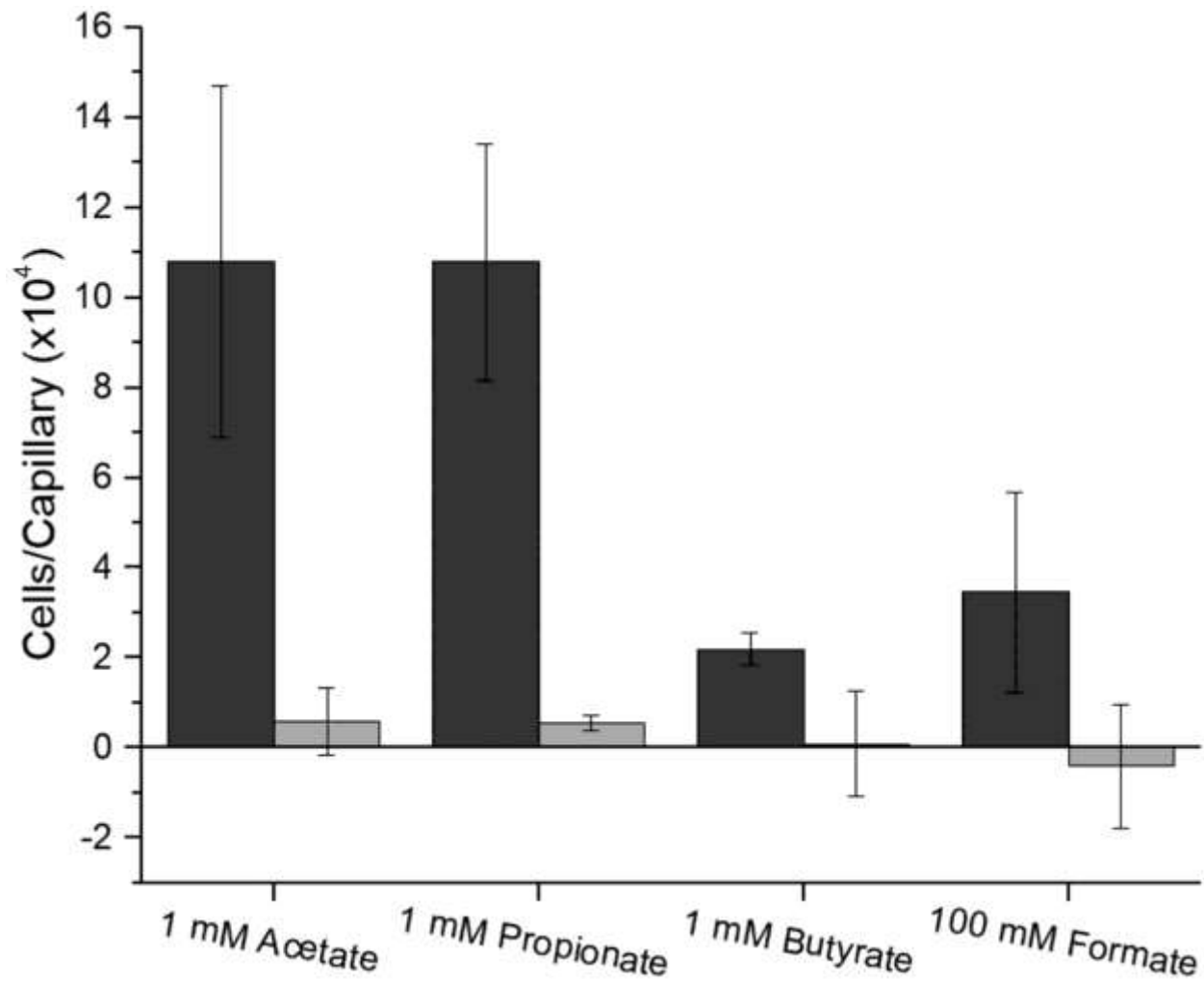




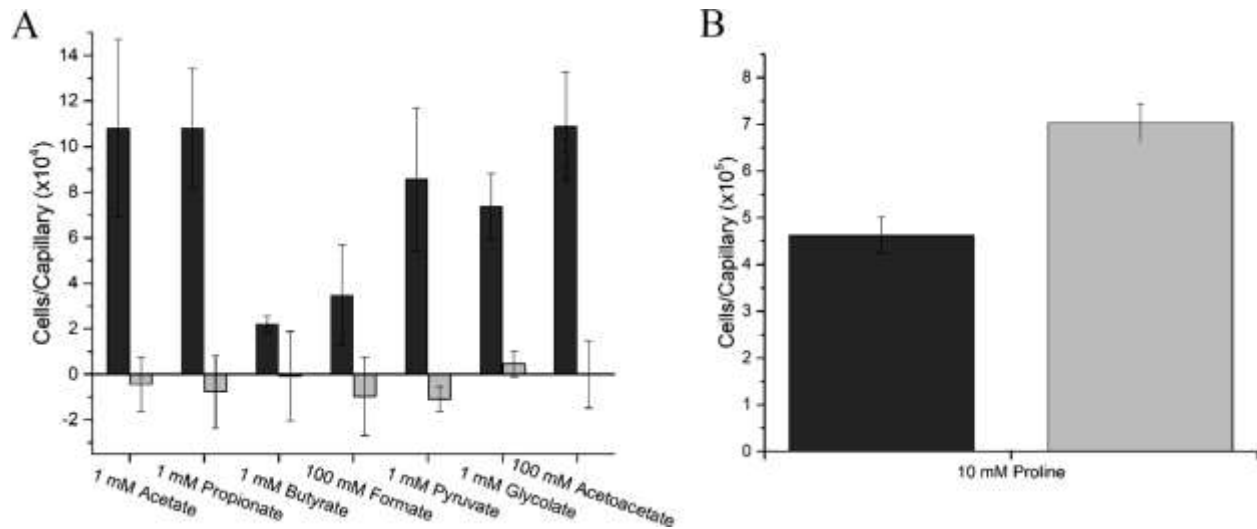
**Fig. 3.2.** High-throughput DSF screen with Biolog plate PM1. The  $\Delta T_m$  is the change in thermal stability of recombinant McpV<sup>PR</sup> in the presence of a compound. The hatched box represents the threshold for a positive interaction. Values above the threshold indicate possible ligand interaction with the protein. Asterisks indicate that no  $T_m$  could be deduced from the melting curves. Values are the mean and standard deviation of three technical replicates.



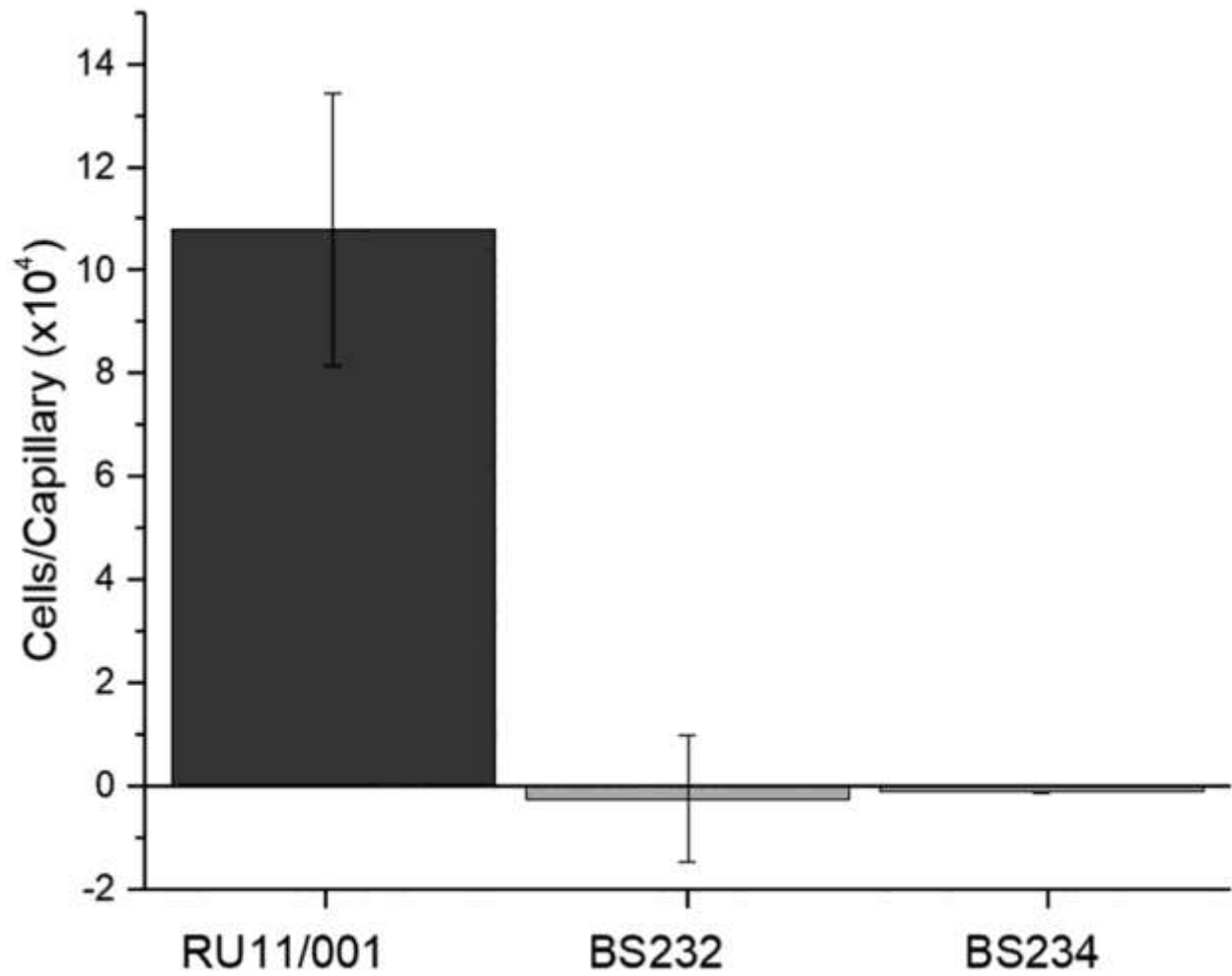
**Fig. 3.3.** *S. meliloti* wild-type chemotaxis responses to carboxylates in the capillary assay. **A.** Dose response curves to acetate (blue), glycolate (cyan), acetoacetate (green), and formate (brown). The last data point of the acetoacetate curve corresponds to a concentration of 250 mM. **B.** Dose response curves to propionate (red), pyruvate (orange), and butyrate (green). The numbers of bacteria accumulated in control capillaries are subtracted from test capillaries to account for random movement of bacteria into capillaries. Values are the mean and standard deviation of three biological replicates.



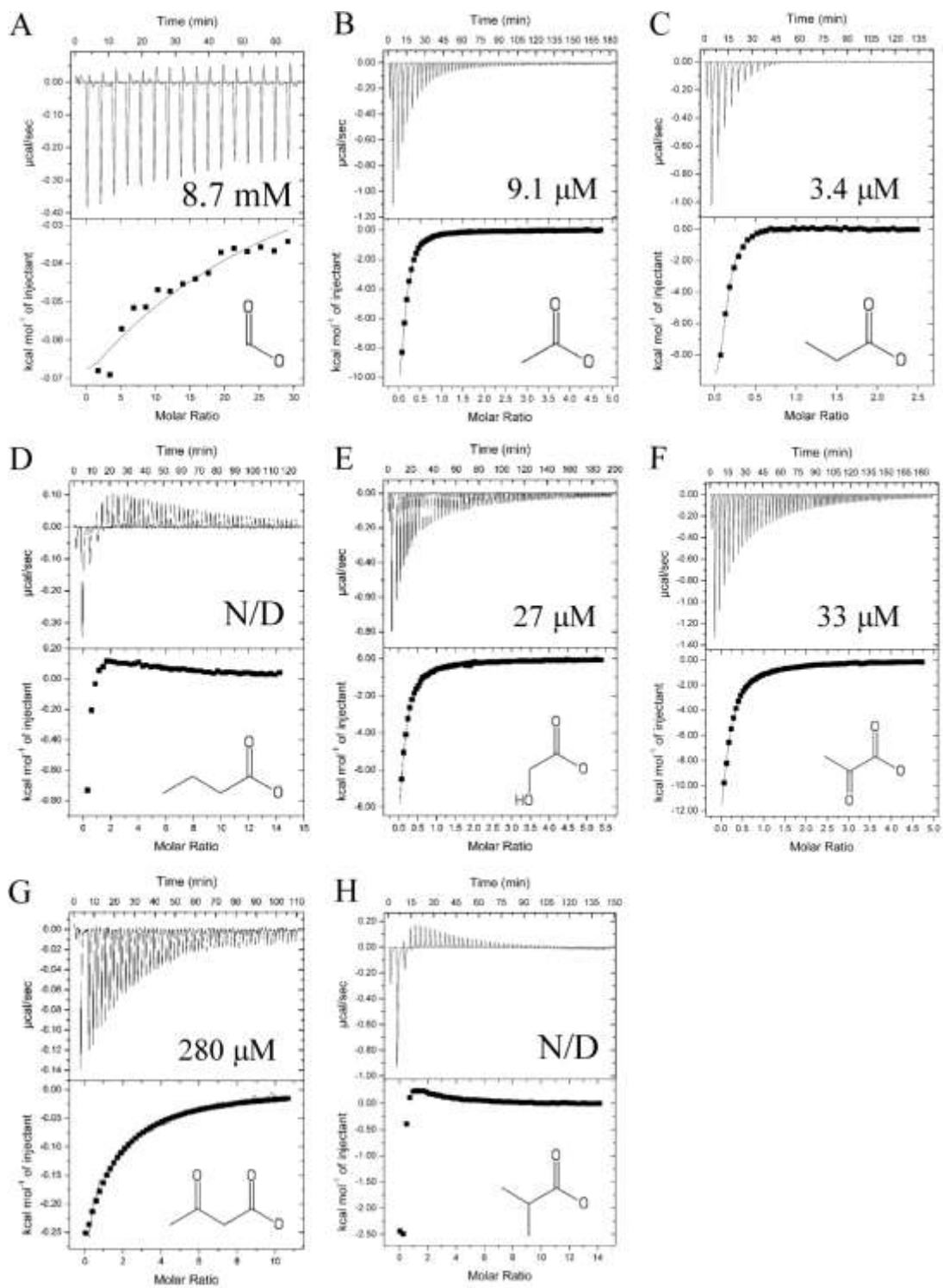
**Fig. 3.4.** Chemotaxis responses of *S. meliloti* wild type (black) and a strain lacking all nine chemoreceptors (*che*, grey) to carboxylates in the capillary assay at peak concentrations of attraction. Chemotaxis data of the wild-type response are taken from Fig. 3.3. Values are the means and standard deviation of three biological replicates.



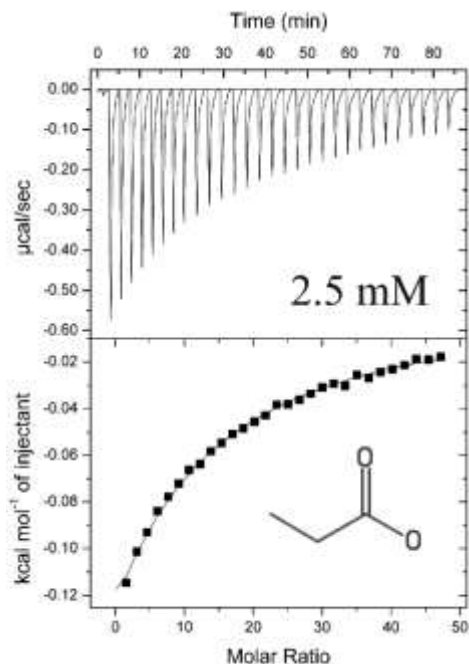
**Fig 3.5.** Chemotaxis responses of *S. meliloti* wild type (black) and a strain lacking *mcpV* (grey) in the capillary assay. **A.** Chemotaxis responses of the wild type and  $\Delta mcpV$  to the peak concentration of acetate, propionate, butyrate, formate, pyruvate, glycolate, and acetoacetate. Chemotaxis data of the wild-type response is taken from Fig. 3.3. **B.** Chemotaxis responses of the wild type and  $\Delta mcpV$  to 10 mM proline. Note the difference in scale between A and B. Values are the means and standard deviation of three biological replicates.



**Fig. 3.6.** *S. meliloti* wild- type, BS232 (McpV<sup>Y143A</sup>), and BS234 (McpV<sup>H103E</sup>) chemotaxis responses to 1 mM propionate in the capillary assay. Chemotaxis data of the wild-type response is taken from Fig. 3.3. Values of the mutant responses are the mean and standard deviation of two biological replicates.



**Fig. 3.7.** Isothermal titration calorimetry of 75  $\mu\text{M}$  recombinant McpV<sup>PR</sup> with carboxylates. The top panels depict the raw titration data and the  $K_d$ . The lower panels are the isotherms derived by integrating peaks from the raw data and the chemical structure of the titrant. **A.** 10 mM formate; **B.** 2 mM acetate; **C.** 2 mM propionate; **D.** 5 mM butyrate; **E.** 2 mM glycolate; **F.** 2 mM pyruvate; **G.** 5 mM acetoacetate; **H.** 5 mM isobutyrate. N/D not determined. Titrations of ligand in to buffer without protein were performed to subtract heats of dilution. Dissociation constants were reported from curves generated using the MicroCal version of Origin 7.0 software using the “one binding sites” model (Origin Lab, Northampton, MA).



**Fig. 3.8.** Isothermal titration calorimetry of 75  $\mu\text{M}$  recombinant  $\text{McpV}^{\text{Y143A-PR}}$  with 15 mM propionate. The top panel depicts the raw titration data and the  $K_d$ . The lower panel is the isotherm derived by integrating peaks from the raw data and the chemical structure of the titrant. Titrations of ligand into buffer without protein were performed to subtract heats of dilution. Dissociation constants were reported from curves generated using the MicroCal version of Origin 7.0 software using the “one binding sites” model. (Origin Lab, Northampton, MA).



**Chapter 4 - McpT is a broad range carboxylate chemoreceptor in *Sinorhizobium meliloti***

HIBA BAAZIZ<sup>1</sup>, K. KARL COMPTON<sup>1</sup>, BIRGIT E. SCHARF<sup>1</sup>

<sup>1</sup>Department of Biological Sciences, Life Sciences I, Virginia Tech, Blacksburg, VA 24061,  
USA

Running title: McpT is a carboxylate sensor

Key words: chemotaxis, legume, rhizobia, symbiosis

\*For correspondence:

E-mail [bscharf@vt.edu](mailto:bscharf@vt.edu)

Tel (+1) 540 231 0757

Fax (+1) 540 231 4043

Biological Sciences, Life Sciences I

Virginia Tech

Blacksburg, VA 24061, USA

**Attribution: HB and KC designed experiments and wrote the manuscript. HB performed experiments for Figs. 4.1 – 4.9. KC performed experiments for Figs. 4.10 – 4.13. BS designed experiments, provided technical guidance, and revisions on the manuscript.**

## ABSTRACT

Chemoreceptors enable the legume symbiont *Sinorhizobium meliloti* to detect and respond to specific chemicals released from their particular host plant alfalfa. The periplasmic regions of chemoreceptors (PR) act as the sensory input modules for chemotaxis systems via binding, either directly or indirectly, of specific ligands. *S. meliloti* has eight chemoreceptors, however, only the function of three have been characterized, with McpU, McpV and McpX serving as general amino acid, short-chain carboxylates, and quaternary ammonium compound sensors, respectively. In the present study, we have characterized *S. meliloti* chemoreceptor McpT. High-throughput differential scanning fluorimetry assays identified fifteen potential ligands for McpT<sup>PR</sup>, the majority classified as mono-, di- and tri-carboxylates. *S. meliloti* exhibits positive chemotaxis toward seven selected carboxylates, namely, alpha-ketobutyrate, citrate, glyoxylate, malate, malonate, oxalate, and succinate. All these carboxylates were detected in alfalfa seed exudates. Deletion of *mcpT* showed a significant decrease of chemotaxis to all carboxylates except for citrate; its sensing was only marginally affected. Isothermal titration calorimetry revealed that McpT<sup>PR</sup> binds specifically to the monocarboxylate glyoxylate as well as the dicarboxylates malonate and oxalate. However, no direct binding was detected for the remaining four carboxylates that elicited an McpT-dependent chemotaxis response. Taken together, these results demonstrate that McpT is a broad range carboxylate chemoreceptor that mediates chemotactic response via direct ligand binding and another mechanism that yet needs to be identified.

## INTRODUCTION

Responding to changes in environmental conditions is a fundamental strategy of bacteria to survive and proliferate (1,2). Bacteria have evolved an advanced sensing mechanism, named chemotaxis, that allows them to rapidly respond to chemical gradients in their surroundings, approaching chemically favorable environments, and avoiding hostile ones (3). Bacterial chemotaxis is implicated in the establishment of a symbiosis between Fabaceae (legumes) and nitrogen-fixing soil bacteria, referred to as rhizobia. Rhizobia sense signaling biomolecules released by germinating host seeds and developing roots, and modulate their swimming direction towards increasing concentrations of these compounds accumulated in the spermosphere and rhizosphere (4).

Chemotaxis of the soil-dwelling  $\alpha$ -proteobacterium *Sinorhizobium meliloti* plays a key role in triggering the colonization of its host legume *Medicago sativa* (alfalfa) roots, ultimately resulting in the development of symbiotic root nodules populated by nitrogen-fixing *S. meliloti* bacteroides (5,6). Prior to nodulation, germinating alfalfa seeds exude a large variety of signaling metabolites including sugars, amino acids, organic acids, and quaternary ammonium compounds that act as chemoattractant for *S. meliloti* (4,7–9). Signal perception is mediated through the sensory domains of chemoreceptors, known as Methyl-accepting Chemotaxis Proteins (MCPs). The direct binding of chemoattractants to the sensory domain, the ligand binding domain (LBD), is the most common mechanism of signal perception by MCPs (10). However, sensing can also be indirect through binding of the ligand to a cognate periplasmic binding proteins (BPs). The ligand-BPs complexes are then able to bind to the LBD and consequently trigger the chemotactic response (11). A well-studied example of indirect ligand sensing is maltose chemotaxis in *Escherichia coli*, where a periplasmic maltose-binding protein binds maltose prior to interaction with the sensory domain of the Tar chemoreceptor (12,13).

Some sensory domains can also contain cofactors, such as heme or flavin adenine dinucleotide (FAD), which allow chemoreceptors to detect oxygen and changes in redox status (14,15). In the intensely studied *E. coli* signaling pathway, signal perception by the LBDs of four transmembrane chemoreceptors (Tar, Tsr, Trg and Aer) generates a molecular stimulus that modulates the autophosphorylation activity of the histidine autokinase, CheA, which in turn leads to phosphorylation of the response regulator CheY. Phosphorylated CheY interacts with the flagellar

motor to control the direction of flagellar rotation and to ultimately mediate chemotaxis (2,15–19). In contrast, *S. meliloti* harbors genes coding for six transmembrane chemoreceptors (McpT, McpU, McpV, McpW, McpX, and McpZ) as well as two cytosolic receptors (McpY and IcpA). To modulate swimming behavior, the signaling domains of transmembrane chemoreceptors convey sensory information to the flagellar motors, controlling the variation in its rotary speed, across a complex signal transduction pathway that includes nine chemotaxis proteins (named CheA, CheW, CheR, CheB, CheS, CheD, CheT, CheY1 and CheY2) (20).

Different structural organizations have been described for transmembrane chemoreceptors, the most common one is typified by the *E. coli* receptors Tar and Tsr (17). These receptors consists of a variable periplasmic region (PR) flanked by two transmembrane helices, followed by a HAMP (histidine kinase, adenylyl cyclase, MCP, and phosphatase) domain and a highly conserved cytoplasmic signaling domain (17,21). We have shown previously that *S. meliloti* chemoreceptors McpU, McpV, and McpX, containing Cache domains (calcium channels and chemotaxis receptors) in their PR, sense plant-derived amino acids, short-chain carboxylates, and quaternary ammonium compounds, respectively, via direct binding (7–9,20). The transmembrane chemoreceptor McpT is 652 aa in length and consists of a periplasmic region of 150 aa, which harbors the ligand binding domain, two transmembrane domains, and a cytoplasmic region consisting of two methylation helices, a signaling domain, and two HAMP domains (Fig 4.1) (20). The characterization of McpT function as well as its mode of attractant recognition is a focus of the present study. We show that McpT is a broad-range sensor for alfalfa-derived carboxylates. We also demonstrate that McpT<sup>PR</sup> can recognize carboxylates through direct binding and additionally through a yet-to be identified indirect binding mechanism. Overall, our study illustrates that alfalfa exuded-carboxylates play a role in the symbiotic signaling between alfalfa and *S. meliloti*, because *S. meliloti* employs at least two chemoreceptors, McpT and McpV, to assure sensing of a wide range of its host-borne carboxylates.

## RESULTS

### **A high-throughput differential scanning fluorimetry assay to screen the putative ligand profile of McpT**

Bacteria have evolved chemotactic behavior for a vast variety of beneficial compounds, such as carbon, nitrogen, or energy sources (22). To test for potential interaction of this type of compounds with McpT, the periplasmic region of McpT (McpT<sup>PR</sup>; McpT<sub>17-166</sub>) was recombinantly expressed and purified by affinity and size exclusion chromatography. A high-throughput *in vitro* differential scanning fluorimetry (DSF) assay using Biolog Phenotype Microarray™ plates as templates was then used to screen for potential ligands of McpT<sup>PR</sup>. Typically, the binding of a ligand stabilizes the protein causing an increase of its melting temperature ( $T_m$ ). The DSF assay using Biolog plates helps rapid screening of ligand binding by measuring the difference in the protein temperature midpoint of unfolding ( $\Delta T_m$ ) in the presence and absence of potential ligands. We screened the recombinant McpT<sup>PR</sup> against the compounds present in Biolog plates PM1 and PM2A, that represent 190 carbon sources (sugars, carboxylates, amino acids and organic acids), as well as PM3B containing 95 nitrogen sources.

In the absence of ligands, McpT<sup>PR</sup> displayed a  $T_m$  of  $33 \pm 0.5$  °C. The melting temperature of the McpT<sup>PR</sup> in the presence of most compounds was within 1 to 2 °C of the control, therefore, a positive interaction was defined as a  $T_m$  shift  $> 2.5$  °C. Of the 258 potential ligands screened, no positive interaction was observed with compounds from plate PM3B and only 15 compounds from PM1 and PM2B caused temperature shifts with  $\Delta T_m$  values  $> 2.5$  °C (Fig. 4.2 and 4.3). Twelve of these compounds were identified as mono-, di- and tri-carboxylates. The monocarboxylates  $\alpha$ -ketobutyrate and glyoxylate elicited the greatest thermal shifts with a  $\Delta T_m$  of 5 °C, followed by propionate that resulted in a  $\Delta T_m$  of 4 °C. The dicarboxylates L-malate, D-malate, and D-, L-malate, as well as the monocarboxylate acetate led to a  $\Delta T_m$  of 3.5 °C. The  $T_m$  of McpT<sup>PR</sup> shifted by 3 °C in the presence of the dicarboxylates tartrate, oxalate, and the tricarboxylate oxalomalate. The lowest shift just at background level ( $\Delta T_m$  of 2.5 °C) was elicited by the dicarboxylate succinate and the tricarboxylate citrate (Fig. 4.2). In addition to carboxylates, three carboxylate derivatives, namely methyl-D-lactate, citramalate, and salicylate exhibited  $\Delta T_m$  values above the threshold of 2.5 °C (Fig. 4.2).

Although DSF screening using Biolog plates provides rapid information on putative ligands, it does not represent definitive proof of binding, as proteins can interact with the minimal medium present in the Biolog wells causing false positive results. Therefore, binding of the twelve carboxylates identified as potential ligands of McpT<sup>PR</sup> was tested with defined ligand solutions.

The DSF assay was performed at three different ligand concentrations (0.4, 4, and 40 mM) to test whether the protein melting temperature increases with fractional ligand occupancy (23). Increasing concentrations of tartrate and propionate did not increase the thermal stability of the protein, indicating that these compounds are not ligands for McpT<sup>PR</sup>. However, increasing concentrations of the remaining carboxylates resulted in a statistically significant increase in the thermal stability of the protein, confirming that they are potential ligands for McpT<sup>PR</sup> (Fig. 4.4). It is noteworthy that the greatest thermal shifts were elicited by 40 mM glyoxylate and oxalomalate with a  $\Delta T_m$  of 6.75°C, and 6.5°C, respectively (Fig. 4.4).

### **Chemotaxis of *S. meliloti* wild type to carboxylates**

The chemotactic response of *S. meliloti* wild type (RU11/001) towards six representative carboxylates, namely  $\alpha$ -ketobutyrate, citrate, glyoxylate, L-malate, oxalate, and succinate, was tested with quantitative capillary assays. In these assays, the response is quantified by comparing the numbers of cells that navigate into a capillary containing a potential chemoattractant, versus the number of cells that accumulate in a capillary containing only RB medium. The chemotactic response was measured over a range of concentrations for each ligand (10  $\mu$ M to 2 M). *S. meliloti* wild type showed positive chemotaxis towards all six compounds (Fig. 4.5). All compounds elicited a concentration-dependent chemoattraction curve that peaked and subsequently decreased, except for succinate and oxalate, where the curve peaked at their maximum solubility limit in RB medium. The wild-type strain was attracted to  $\alpha$ -ketobutyrate, glyoxylate, oxalate, and succinate with a chemoattraction peak at 10 mM, 100 mM, 250 mM and 400 mM, respectively. Malate elicited a large chemoattraction plateau between 100 mM and 1 M. Chemoattraction to citrate was observed at lower concentrations, which peaked at 100  $\mu$ M. At peak concentration, oxalate and glyoxylate caused the greatest accumulation with 61,000 and 52,000 cells per capillary, respectively. Succinate and malate followed each with 35,000 and 30,000 cells per capillary. The peak concentration of  $\alpha$ -ketobutyrate elicited the accumulation of 23,300 cells per capillary. Citrate displayed the lowest accumulation with only 14,000 cells per capillary (Fig. 4.5).

### **McpT mediates carboxylate sensing in *S. meliloti***

To investigate the impact of McpT on carboxylate sensing, the chemotaxis response of a strain lacking *mcpT* (RU11/838) was tested toward the six selected carboxylates at different concentrations. In the absence of *mcpT*, *S. meliloti* chemotaxis to carboxylates was significantly

decreased (Fig. 4.6). For succinate, the number of  $\Delta mcpT$  cells accumulated at the peak concentration (400 mM) was 6-fold lower compared to that of the wild type. Moreover, chemoattraction of the  $\Delta mcpT$  strain to the peak concentrations of malate (1 M) and oxalate (250 mM) was also decreased by 4-fold, while attraction to the peak concentration of  $\alpha$ -ketobutyrate (10 mM) and glyoxylate (100 mM) was lower by 3-fold. It is noteworthy that in the absence of *mcpT*, the peak response of  $\alpha$ -ketobutyrate was shifted to a higher concentration (100 mM), which may indicate that cells became less sensitive. In addition, chemotaxis toward citrate was marginally reduced compared to the other carboxylates, since the strain lacking *mcpT* exhibited a chemoattraction decrease of 1.5-fold at the peak concentration (0.1 mM) (Fig. 4.6). Since McpV is also a sensor of carboxylates, we performed chemotaxis assays with BS275, a strain lacking both *mcpT* and *mcpV*. The double deletion  $\Delta mcpT \Delta mcpV$  strain displayed a severe decrease of chemotaxis to the peak concentration of succinate by 6-fold. The absence of both MCPs completely abolished  $\alpha$ -ketobutyrate and citrate sensing (Fig 4.7). However, the chemoattraction of the double deletion strain to the peak concentration of malate was similar to that of  $\Delta mcpT$  strain (Fig 4.8). It should be noted that the absence of *mcpT* as well as the absence of both MCPs ( $\Delta mcpT \Delta mcpV$ ) does not negatively impact the chemotaxis ability to an attractant sensed by other chemoreceptors. The deletion strains  $\Delta mcpT$  and  $\Delta mcpT \Delta mcpV$  reacted to proline, which is sensed by McpU and McpX, with the same strength as the wild-type strain (Fig. 4.9). The introduction of the complementing plasmid pBBR1MCS-2 constitutively expressing *mcpT* (pBS1055) to the  $\Delta mcpT$  strain (RU11/838) restored the chemotaxis toward carboxylates to wild-type levels (Fig. 4.7 and 4.8). To attribute the observed bacterial accumulation in the capillaries to chemotaxis, a strain lacking the entire *cheI* operon ( $\Delta cheI$ ; BS251) was tested and exhibited chemoattraction towards malate at levels similar to the  $\Delta mcpT$  strain (Fig. 4.5). In addition, the chemoattraction of a strain lacking all nine chemoreceptors (*che*<sup>-</sup>, RU13/149), towards these citrate, malate, succinate and  $\alpha$ -ketobutyrate was also tested at concentrations of peak attraction. Chemotaxis to succinate, citrate and  $\alpha$ -ketobutyrate was severely diminished (Fig. 4.7). Surprisingly, chemotaxis toward the peak concentration of malate was not fully abolished in the absence of all chemoreceptors, as chemoattraction was observed at levels similar to those of  $\Delta mcpT$  strain (around 5,000 cells per capillary). Together, these results show that *S. meliloti* McpT chemoreceptor senses a broad range of carboxylates.

## **Isothermal titration calorimetry demonstrates direct binding of specific carboxylates to McpT<sup>PR</sup>**

To appropriately verify chemoattractants that directly bind to McpT<sup>PR</sup>, we used Isothermal Titration Calorimetry (ITC). No binding was detected for any of the 4- and 6-C attractants, such as citrate, malate, and succinate (Fig. 4.10). Oxalate was tested next because it is the simplest dicarboxylate that reacted in DSF. Here, we detected binding with a  $K_d$  of 44  $\mu\text{M}$  from an endothermic reaction (Fig. 4.10). The next step was to assay compounds that are chemically similar to oxalate. We tested glyoxylate, oxalaoacetate, and glycolate, but only found glyoxylate to interact with McpT<sup>PR</sup> weakly, with a 370  $\mu\text{M}$   $K_d$  also in an exothermic fashion (Fig. 4.10). Since the buffer may have been too complex and may interfere with binding, we repeated the experiments in a simplified buffer with lower ionic strength, omitting glycerol, and with a neutral pH (0.1 M PIPES, 0.1 M KCl [pH 7.0]). Titrations with citrate, malate,  $\alpha$ -ketobutyrate, and succinate again did not result in binding (Fig. 4.11). However, oxalate interacted with McpT<sup>PR</sup> in this buffer, but the  $K_d$  was measured at 262 nM, which is about 150-fold lower than that in the previous buffer. It should be noted that the McpT<sup>PR</sup> – oxalate interaction in the neutral pH buffer was still exothermic, with an enthalpy change of -7,133 cal/mol and an entropy change of 4.92 cal/mol/deg K. Strangely, glyoxylate did not exhibit any binding in this neutral pH buffer (Fig. 4.11). Finally, when titrating McpT<sup>PR</sup> against malonate, we observed an initial exothermic peak followed by several endothermic peaks (Fig. 4.11). This suggests that more than one interaction was occurring, thus prohibiting the quantification of the McpT<sup>PR</sup> – malonate interaction. We have observed this kind of obscure interaction in previous chemoreceptor studies (8). As a whole, ITC binding experiments suggest that McpT<sup>PR</sup> only interacts with 2- and 3-C dicarboxylates and possibly with glyoxylate.

## **Analytic size exclusion chromatography suggests the formation of McpT<sup>PR</sup> dimers**

A structure-based homology search using the SWISS-MODEL repository predicted that the McpT periplasmic region classifies as a four-helix bundle (Fig 4.1). This type of fold is an obligate dimer that binds one ligand per MCP dimer. To support this prediction, we performed analytical size exclusion chromatography with McpT<sup>PR</sup> and compared its retention time (RT) to that of several molecular weight standards. McpT<sup>PR</sup> eluted at 15.8 ml, just after ovalbumin (RT = 15.4 ml), which has a molecular weight (MW) of 44 kDa. Monomeric McpT<sup>PR</sup> has a molecular weight of 18.3 kDa, suggesting that the McpT<sup>PR</sup> population exists entirely as a dimer (Fig. 4.12).



We hypothesized that the presence of a putative ligand might affect size or multimerization of the protein. Thus, analytical size exclusion experiments were repeated with buffers amended with 5 mM of the ligands tested in ITC. The RT of the protein in the absence of ligand was the highest with 15.79 ml. The RT in the presence of malate was the next highest, 15.74. Experiments with succinate and  $\alpha$ -ketobutyrate had similar RTs at 15.68 and 15.70 ml, respectively. The RT in the presence of citrate, oxalate, and malonate all clustered closely between 15.64 and 15.61 ml (Fig. 4.13). These data show a correlation between RT and binding in ITC whereby the protein with ligand shows a slightly lower retention time than with non-ligands.

## DISCUSSION

Bacteria use chemotaxis to survive in the soil and compete for niches such as root nodules (36-38). Plants secrete chemical cues that attract and support a plethora of microorganisms. Root exudates such as carboxylates have a major impact in shaping the rhizosphere ecology because they modify the soil environment and provide resident biota with carbon sources (39-45). Recruitment of bacteria can occur when the cell follows a chemical gradient of plant exudates using specific chemoreceptors. Here, we describe McpT and its contribution to taxis to a range of carboxylates.

We first elucidated the ligand profile of McpT with a high-throughput *in-vitro* screen. Purified McpT<sup>PR</sup> appeared to interact with a wide range mono-, di-, and, tri-carboxylates. This is a somewhat diverse list of compounds, but nonetheless indicates that McpT interacts with negatively charged organic compounds (Fig 4.2). This finding compelled *in-vivo* behavioral experiments to determine the physiological response to these carboxylates. The strongest attractant in view of numbers of cells drawn to the capillary was oxalate at its maximum test concentration of 250 mM, closely followed in strength by glyoxylate. Malate and succinate were the next strongest and approximately equivalent attractants.  $\alpha$ -ketobutyrate and citrate drew the fewest cells but had lower peak concentrations of attraction (Fig 4.5). Deletion of *mcpT* caused a reduction in chemotaxis for all compounds tested. Furthermore, the wild-type response to malate, succinate,  $\alpha$ -ketobutyrate, and citrate could be rescued by complementation (Fig 4.7 and 4.8). Direct binding studies using ITC were performed in two different buffer systems to gain a deeper understanding

of the protein-ligand interaction in this system. Confoundingly, only glyoxylate, malonate, and oxalate were found to directly interact with McpT<sup>PR</sup>, while the other carboxylates did not. Glyoxylate bound weakly to McpT<sup>PR</sup> and only in one of the tested buffers, namely with a pH of 8.0 and salts and glycerol. Binding to oxalate in the buffer with omitted salts and glycerol and a pH of 7.0 was 150 times stronger than binding in the more complex pH 8.0 buffer. Two explanations seem likely for this disparity: 1) an excess of glycerol and salts may have occluded the binding pocket, making oxalate more difficult to interact with McpT<sup>PR</sup>, or 2) the elevated pH resulted in a more negatively charged protein, which increased electrostatic repulsion between the protein and ligand. Unfortunately, neither option explains the disappearance of binding to glyoxylate in the pH 7.0 buffer (Fig 4.10 and 4.11). Homology modelling suggests the McpT<sup>PR</sup> is a 4-helix bundle domain fold, which forms obligate dimers, and this prediction is supported by analytical size exclusion chromatography resulting in the elution of McpT<sup>PR</sup> at a volume corresponding to approximately twice its molecular weight (Fig 4.1 and 4.12). Analytical size exclusion chromatography in the presence of attractant molecules exhibits a slight shift to earlier RTs. The greatest RT shifts occur in the presence of compounds that were shown via ITC to directly bind to McpT<sup>PR</sup>. This observation could indicate that the carboxylates induce a swelling or bulging of the dimer, thus slightly increasing the size of the complex, which manifests as a reduced RT (Fig 4.13). Perhaps a slightly enlarged protein complex induced by carboxylates may also explain the  $\Delta T_m$  observed in DSF experiments. We would be remiss not to note that slight changes in RT of proteins may also be an artifact from different solute molecules in the chromatography buffer.

Reconciling *in-vitro* binding parameters with physiological responses grants a complete picture of how the bacterium achieves taxis to its attractants. Taxis to glyoxylate and oxalate alone are explained by the canonical direct binding seen in most MCPs (22). In contrast, malate, citrate,  $\alpha$ -ketobutyrate, and succinate do not directly bind to McpT. However, McpT is still implicated in the sensing of these carboxylates (Fig 4.6 and 4.8). The most common explanation is the phenomenon of indirect binding, as described for *E. coli* taxis to maltose and *Bacillus subtilis* taxis to lysine, glutamine, arginine, and methionine (12,13,46). Strains lacking *mcpT* still exhibit taxis to many of these compounds at about 20 % of the wild-type response (Fig 4.7). The remaining response could be due to weak interaction with other MCPs, or the strategy of energy taxis, in which bacteria swim to areas rich in compounds that support redox reactions (47).

This work establishes the involvement of McpT in the sensing of various carboxylates. These molecules are ubiquitous in plant exudates and may be an important avenue of host sensing. Citrate is commonly released by plants to deal with aluminum stress (43). Manganese can be solubilized and made bioavailable by acidification with organic acids like succinate and malonate (48). Succinate and citrate make up most of the carboxylates released from tomato and cucumber roots and are more abundant than exuded sugars (49). Oxalate is released because it can chelate toxic metals like cadmium (50). It is also a chemotactic signal for *Azorhizobium caulinodans* and a critical signal for host plants colonization by plant-beneficial *Burkholderia* spp. (51,52). McpT ligands such as citrate, succinate and malate have also been found in the exudate of germinating alfalfa seedlings (unpublished data). Most carboxylates can be utilized as carbon energy sources via the citric acid cycle; moreover, succinate is a preferred carbon source (53,54). Having sensors spurring navigation to carbon sources and host signals such as the carboxylates is clearly advantageous for soil bacteria such as *S. meliloti*. Carboxylates are a very broad class of compound but, prior to this work, only 2 – 4 C monocarboxylates were known to be attractants for *S. meliloti* (8). This present study expands our knowledge of the sensory repertoire of this bacterium and adds to our model of plant host sensing.

## MATERIAL AND METHODS

### Strains and plasmids

Derivatives *Escherichia coli* K-12 strains, the highly motile derivatives of *S. meliloti* MV II-1 (24), and the plasmids used are listed in Table S1.

### Media and growth conditions

Lysogeny broth (LB) was used to grow *E. coli* strains at 37 °C (25). Tryptone Yeast-Extract Calcium chloride (TYC) medium supplemented with streptomycin (600 µg/ml) was used to grow *S. meliloti* strains at 30 °C (26). Rhizobium basal medium (RB) [6.1 mM K<sub>2</sub>HPO<sub>4</sub>, 3.9 mM KH<sub>2</sub>PO<sub>4</sub>, 1 mM MgSO<sub>4</sub>, 1 mM (NH<sub>4</sub>)<sub>2</sub>SO<sub>4</sub>, 0.1 mM CaCl<sub>2</sub>, 0.1 mM NaCl, 0.01 mM Na<sub>2</sub>MoO<sub>4</sub>, 0.001 mM FeSO<sub>4</sub>, 20 µg of biotin/liter, 100 µg of thiamine/liter] (27), layered on Bromfield agar plates at 30 °C was used to grow motile cells of *S. meliloti* strains for capillary assays (28).

Ampicillin and kanamycin were used for *E. coli* at the final concentrations of 100 µg/ml and 25 µg/ml, respectively. For *S. meliloti*, neomycin and streptomycin were at the final concentrations of 120 µg/ml and 600 µg/ml, respectively.

### **Expression and purification of McpT<sup>PR</sup>**

The recombinant, ligand-binding, periplasmic region of McpT (McpT<sup>PR</sup>; McpT<sub>17-166</sub>) was overproduced from plasmid pTYB11 in *E. coli* ER2566, providing an intein-chitin binding domain (intein-CBD) tag (Table 1S). Four liters of cell culture were grown to an OD<sub>600</sub> of 0.8 at 37 °C in LB containing 100 µg of ampicillin/ml and gene expression was induced by 0.6 mM isopropyl-β-D-thiogalactopyranoside (IPTG) for 16 hours at 16 °C until harvest. Cell pellets were suspended in column buffer (20 mM Tris/HCl, 500 mM NaCl, 1 mM EDTA and 10 % glycerol [pH 8.0]) supplemented with 10 mg/ml of DNase, 1 mM PMSF and 1 × Halt™ Protease Inhibitor Cocktail (Thermo Fischer Scientific). Three passages through a French pressure cell at 16,000 lbs/in<sup>2</sup> (SLM Aminco, Silver Spring, MD) were performed to lyse the cells before clearing the lysate by centrifugation at 56,000 × *g* for 1 hour at 4 °C. Supernatant fractions were purified through 30 ml of settled chitin agarose (New England BioLabs) in a column with 2.5 cm previously equilibrated with column buffer. The intein-CBD-tag cleavage was performed using a cleavage buffer (column buffer supplemented with 50 mM DTT) prior to 48 hours incubation at 4 °C. Proteins were eluted with column buffer and concentrated using an Amicon ultrafiltration system with regenerated cellulose membranes (10-kDa cutoff) (Millipore, Billerica, MA), and further purified by fast protein liquid chromatography on a size exclusion HiPrep 26/60 Sephacryl S-200 HR column (GE Healthcare Life Sciences). The column was equilibrated in protein buffer (100 mM Tricine, 150 mM NaCl, 1 mM EDTA, 15 % glycerol [pH 8.0]) and separation was performed at a flow rate of 1 ml/min. Protein-containing fractions were then concentrated by ultrafiltration on regenerated cellulose membranes. Protein concentration was determined using the Bradford protein assay (BIO-RAD).

### **Analytical size exclusion chromatography**

Experiments utilized a Superdex 200 Increase 10/300 GL column operated by an ÄKTA pure FPLC unit (GE Healthcare, Chicago IL). The injection volume for all samples was 200 µl and the column was developed at a flow rate of 0.5 ml/min. A calibration curve was created using 0.1 mg blue dextran (MW approx. 2,000 kDa), 0.2 mg aldolase (MW 158 kDa), 0.2 mg conalbumin (MW

75 kDa), 0.2 mg ovalbumin (MW 44 kDa), 0.04 mg cytochrome C (MW 12.4 kDa), and 0.02 mg vitamin B12 (MW 1.4 kDa). Excluding the native sample, all McpT<sup>PR</sup> samples were taken from titration experiments and diluted to 30  $\mu$ M (corresponding to about 0.11 mg of injected protein). The chromatography buffer was 0.1 M PIPES, 0.1 M KCl [pH 7.0] for the calibration standards and the native protein. The buffer used in experiments with titrated McpT<sup>PR</sup> was amended with the respective titrant to 5 mM.

### **Differential scanning fluorimetry**

Putative McpT<sup>PR</sup> ligands were investigated by screening Biolog Phenotype MicroArray<sup>TM</sup> (PM) compounds supplied in a 96-well microplate format (BIOLOG, Inc., Hayward, CA). For initial high-throughput screening, compounds in PM1, PM2 and PM3B microplates were dissolved in 35  $\mu$ l of a master mix containing 40  $\mu$ M McpT<sup>PR</sup> and 2 $\times$  Sypro Orange (Invitrogen, Grand Island, NY) in the protein buffer. A volume of 30  $\mu$ l from each well was transferred to a 96-well plate for use in an ABI 7300 real-time PCR system. Thermal denaturation was carried out by increasing the temperature from 10 to 95  $^{\circ}$ C with a 30-s equilibration at each half degree Celsius. The melting temperature ( $T_m$ ) of the protein in each well corresponds to the maximum value of the first derivative of the fluorescence curve. The melting temperature shift ( $\Delta T_m$ ) was determined by subtracting the  $T_m$  of the control well containing no ligand from the  $T_m$  of each test well. The screen was performed in three biological replicates using three Biolog plates. The compounds that yielded a significant ( $> 2.5$   $^{\circ}$ C) positive shift in  $T_m$  were taken as potential ligands. Initial hits were further tested using known final concentrations (0.4, 4.0, and 40 mM) of each potential ligand to confirm binding. Ligands were prepared in protein buffer and the experiments were conducted as described above.

### **Isothermal titration calorimetry**

Direct binding studies were performed using a VP-ITC microcalorimeter (Malvern, United Kingdom). All ligand solutions were prepared in the buffer that was eluted from the preparative size exclusion chromatography column. Titrations were all performed at 10  $^{\circ}$ C, stirring speed 220 rpm, reference power 25  $\mu$ Cal/sec, and with protein in the sample cell. The initial buffer consisted of 0.1 M Tricine, 0.2 M NaCl, 0.2 M KCl, 15 % glycerol [pH 8.0]. The protein concentration was 40  $\mu$ M, except for the malate titration, which used 70  $\mu$ M McpT<sup>PR</sup>. All ligands were used at a

concentration of 5 mM. Titrant was added in 28 ten  $\mu\text{l}$  injections at 0.5  $\mu\text{l}/\text{sec}$  after an initial 1  $\mu\text{l}$  injection.

In the subsequent set of experiments, McpT<sup>PR</sup> was purified and tested in a buffer containing 0.1 M PIPES, 0.1 M KCl [pH 7.0]. Protein was used at a concentration of 45  $\mu\text{M}$ . Citrate, malate,  $\alpha$  - ketobutyrate, succinate, and malonate were titrated at 2 mM, glyoxylate at 1 mM, and oxalate at 250  $\mu\text{M}$ . The initial injection was 0.5  $\mu\text{l}$  followed by 32 seven  $\mu\text{l}$  injections. All data were fit using the VP ITC version of Origin 7 (Origin Labs, Northampton, MA) and the “one-binding site” model.

### **Capillary assays**

Traditional Adler capillary assays (29) with the modification described previously by Compton *et al.* (8) were performed. Briefly, *S. meliloti* cells were grown in RB overlain onto Bromfield agar plates for 16 h at 30 °C. Cells were gently harvested between an OD<sub>600</sub> of 0.15 and 0.17 before being suspended in RB medium to a final OD<sub>600</sub> of 0.15. Three-hundred and fifty microliter of motile *S. meliloti* cells were injected into a pond formed from a U-shaped glass tube between two glass plates. Microcap glass 1  $\mu\text{L}$  capillaries (Drummond Microcaps) flame-sealed at one end were placed into various dilutions of compound solution in a vacuum chamber to fill capillaries. Capillaries were placed into the bacterial ponds and left to incubate at room temperature for two hours. After incubation, the contents of the capillaries were expelled into 1 ml RB medium. Serial dilutions were plated on TYC plates supplemented with streptomycin and subsequent counting of colony forming units (CFUs) were performed. The counts of a control capillary, containing only RB medium, were subtracted from all test capillaries. Accumulation of bacterial cells in the capillaries was calculated as the average from the CFUs obtained in triplicate plates and the results were expressed as the mean of at least four separate capillary assays for each compound and concentration. The relative chemotaxis response was calculated as the ratio of the accumulation of the deletion mutant in the capillaries to that of the wild type.

### **ACKNOWLEDGMENTS**

This study was supported by NSF grants MCB-1253234 and MCB-1817652 to Birgit Scharf.

**Table 1. Bacterial strains and plasmids**

Species, strain, or plasmid	Characteristics <sup>a</sup>	Reference or source
<b>Strains</b>		
<i>E. coli</i>		
DH5 $\alpha$	<i>recA1 endA1</i>	(30)
ER2566	<i>lon ompT lacZ::T7</i>	New England Biolabs
S17-1	<i>recA endA thi hsdR</i> RP4-2 Tc::Mu::Tn7 Tp <sup>r</sup> Sm <sup>r</sup>	(31)
<i>S. meliloti</i>		
BS251	Sm <sup>r</sup> ; $\Delta icpA$ , $\Delta cheS$ , $\Delta cheY1$ , $\Delta cheA$ , $\Delta cheW$ , $\Delta cheR$ , $\Delta cheB$ , $\Delta cheY2$ , $\Delta cheD$ , and $\Delta cheT$ . ( $\Delta cheI$ )	This work
BS275	Sm <sup>r</sup> ; $\Delta mcpT$ $\Delta mcpV$	This work
RU11/001	Sm <sup>r</sup> ; spontaneous streptomycin-resistant wild-type strain	(32)
RU11/838	Sm <sup>r</sup> ; $\Delta mcpT$	(20)
RU13/149	Sm <sup>r</sup> ; $\Delta mcpS$ $\Delta mcpT$ $\Delta mcpU$ $\Delta mcpV$ $\Delta mcpW$ $\Delta mcpX$ $\Delta mcpY$ $\Delta mcpZ$ $\Delta icpA$ ( $\Delta 9$ )	(20)
RU13/340	Sm <sup>r</sup> and Neo <sup>r</sup> ; $\Delta mcpT$ containing pBS1055 ( $\Delta mcpT/mcpT$ )	This work
<b>Plasmids</b>		
pBBR1MCS-2	Km <sup>r</sup> ; expression vector	(33)
pBS425	Ap <sup>r</sup> ; SapI/SpeI PCR fragment containing <i>mcpT</i> , 450 bp (150 aa), cloned into pTYB11	This work

pBS1055	Km <sup>r</sup> ; HindIII/XbaI PCR fragment containing <i>mcpT</i> 48- 498 bp (McpT <sup>PR</sup> , aa 17-166) cloned into pBBR1MCS-2	This work
pTYB11	Ap <sup>r</sup> ; expression vector	New England Biolabs

<sup>a</sup>Nomenclature is presented according to Bachmann (34) and Novick *et al.* (35). Tp<sup>r</sup>, trimethoprim resistance; Ap<sup>r</sup>, ampicillin resistance; Sm<sup>r</sup>, streptomycin resistance; Km<sup>r</sup>, kanamycin resistance.



## REFERENCES

1. Raina J-B, Fernandez V, Lambert B, Stocker R, Seymour JR. The role of microbial motility and chemotaxis in symbiosis. *Nat Rev Microbiol*. 2019 May 1;17(5):284–94.
2. Hazelbauer GL, Falke JJ, Parkinson JS. Bacterial chemoreceptors: high-performance signaling in networked arrays. *Trends Biochem Sci*. 2007/12/31 ed. 2008 Jan;33(1):9–19.
3. Hoch JA. Two-component and phosphorelay signal transduction. *Curr Opin Microbiol*. 2000 Apr 1;3(2):165–70.
4. Nelson EB. Microbial dynamics and interactions in the spermosphere. *Annu Rev Phytopathol*. 2004;42:271–309.
5. Caetano-Anollés G, Wall LG, De Micheli AT, Macchi EM, Bauer WD, Favelukes G. Role of motility and chemotaxis in efficiency of nodulation by *Rhizobium meliloti*. *Plant Physiol*. 1988 Apr;86(4):1228–35.
6. Jones KM, Kobayashi H, Davies BW, Taga ME, Walker GC. How rhizobial symbionts invade plants: the *Sinorhizobium*–*Medicago* model. *Nat Rev Microbiol*. 2007 Aug 1;5(8):619–33.
7. Webb BA, Hildreth S, Helm RF, Scharf BE. *Sinorhizobium meliloti* chemoreceptor McpU mediates chemotaxis toward host plant exudates through direct proline sensing. *Appl Environ Microbiol*. 2014 Jun;80(11):3404–15.
8. Compton KK, Hildreth SB, Helm RF, Scharf BE. *Sinorhizobium meliloti* Chemoreceptor McpV senses short-chain carboxylates via Direct Binding. Stock AM, editor. *J Bacteriol*. 2018 Dec 1;200(23):e00519-18.
9. Webb BA, Karl Compton K, Castañeda Saldaña R, Arapov TD, Keith Ray W, Helm RF, et al. *Sinorhizobium meliloti* chemotaxis to quaternary ammonium compounds is mediated by the chemoreceptor McpX. *Mol Microbiol*. 2017 Jan 1;103(2):333–46.
10. Mowbray SL, Koshland DE. Mutations in the aspartate receptor of *Escherichia coli* which affect aspartate binding. *J Biol Chem*. 1990 Sep 15;265(26):15638–43.
11. Gardina PJ, Bormans AF, Hawkins MA, Meeker JW, Manson MD. Maltose-binding protein interacts simultaneously and asymmetrically with both subunits of the Tar chemoreceptor. *Mol Microbiol*. 1997 Mar 1;23(6):1181–91.
12. Hazelbauer GL. Maltose chemoreceptor of *Escherichia coli*. *J Bacteriol*. 1975 Apr 1;122(1):206.
13. Kossmann M, Wolff C, Manson MD. Maltose chemoreceptor of *Escherichia coli*: interaction of maltose-binding protein and the tar signal transducer. *J Bacteriol*. 1988 Oct 1;170(10):4516.
14. Hou S, Larsen RW, Boudko D, Riley CW, Karatan E, Zimmer M, et al. Myoglobin-like aerotaxis transducers in Archaea and Bacteria. *Nature*. 2000 Feb 1;403(6769):540–4.
15. Bibikov SI, Biran R, Rudd KE, Parkinson JS. A signal transducer for aerotaxis in *Escherichia coli*. *J Bacteriol*. 1997 Jun;179(12):4075–9.

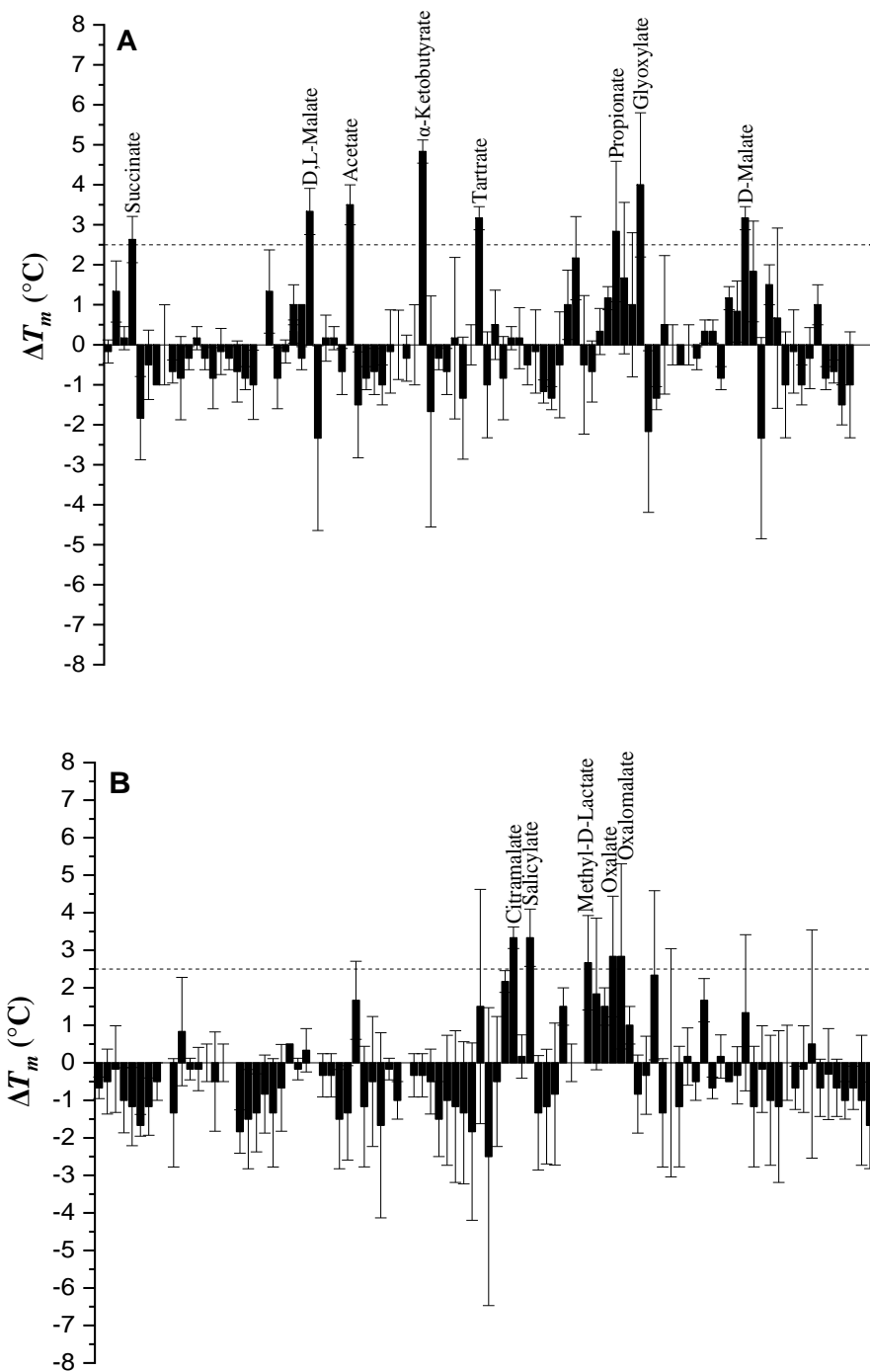
16. Falke JJ, Bass RB, Butler SL, Chervitz SA, Danielson MA. The two-component signaling pathway of bacterial chemotaxis: A Molecular View of Signal Transduction by Receptors, Kinases, and Adaptation Enzymes. *Annu Rev Cell Dev Biol.* 1997 Nov 1;13(1):457–512.
17. Wadhams GH, Armitage JP. Making sense of it all: bacterial chemotaxis. *Nat Rev Mol Cell Biol.* 2004 Dec 1;5(12):1024–37.
18. Sourjik V, Wingreen NS. Responding to chemical gradients: bacterial chemotaxis. *Cell Regul.* 2012 Apr 1;24(2):262–8.
19. Parkinson JS, Hazelbauer GL, Falke JJ. Signaling and sensory adaptation in *Escherichia coli* chemoreceptors: 2015 update. *Trends Microbiol.* 2015/03/30 ed. 2015 May;23(5):257–66.
20. Meier VM, Muschler P, Scharf BE. Functional analysis of nine putative chemoreceptor proteins in *Sinorhizobium meliloti*. *J Bacteriol.* 2007 Mar 1;189(5):1816.
21. Briegel A, Ortega DR, Tocheva EI, Wuichet K, Li Z, Chen S, et al. Universal architecture of bacterial chemoreceptor arrays. *Proc Natl Acad Sci.* 2009 Oct 6;106(40):17181.
22. Ortega Á, Zhulin IB, Krell T. Sensory repertoire of bacterial chemoreceptors. *Microbiol Mol Biol Rev.* 2017 Dec 1;81(4):e00033-17.
23. Vivoli M, Novak HR, Littlechild JA, Harmer NJ. Determination of protein-ligand interactions using differential scanning fluorimetry. *J Vis Exp JoVE.* 2014 Sep 13;(91):51809–51809.
24. Kamberger W. An Ouchterlony double diffusion study on the interaction between legume lectins and rhizobial cell surface antigens. *Arch Microbiol.* 1979 Apr 1;121(1):83–90.
25. Bertani G. Studies on lysogenesis I. *J Bacteriol.* 1951 Sep 1;62(3):293.
26. Scharf B, Schuster-Wolff-Bühring H, Rachel R, Schmitt R. Mutational Analysis of the *Rhizobium lupini* H13-3 and *Sinorhizobium meliloti* flagellin genes: Importance of flagellin a for flagellar filament structure and transcriptional regulation. *J Bacteriol.* 2001 Sep 15;183(18):5334.
27. Götz R, Limmer N, Ober K, Schmitt R. Motility and chemotaxis in two strains of *Rhizobium* with complex flagella. *Microbiology.* 1982;128(4):789–98.
28. Sourjik V, Schmitt R. Different roles of CheY1 and CheY2 in the chemotaxis of *Rhizobium meliloti*. *Mol Microbiol.* 1996 Nov;22(3):427–36.
29. Adler J. A method for measuring chemotaxis and use of the method to determine optimum conditions for chemotaxis by *Escherichia coli*. *J Gen Microbiol.* 1973 Jan 1;74(1):77–91.
30. Hanahan D, Meselson M. Plasmid screening at high colony density. *Methods Enzymol.* 1983;100:333–42.
31. Simon R, O’Connell M, Labes M, Pühler A. Plasmid vectors for the genetic analysis and manipulation of rhizobia and other gram-negative bacteria. *Methods Enzymol.* 1986;118:640–59.
32. Pleier E, Schmitt R. Expression of two *Rhizobium meliloti* flagellin genes and their contribution to the complex filament structure. *J Bacteriol.* 1991 Mar;173(6):2077–85.

33. Kovach ME, Elzer PH, Steven Hill D, Robertson GT, Farris MA, Roop RM, et al. Four new derivatives of the broad-host-range cloning vector pBBR1MCS, carrying different antibiotic-resistance cassettes. *Gene*. 1995 Dec 1;166(1):175–6.
34. Bachmann B. Linkage map of *Escherichia coli* K-12, edition 8. *Microbiol Rev*. 1990 Jun;54(2):130–97.
35. Novick RP, Clowes RC, Cohen SN, Curtiss R 3rd, Datta N, Falkow S. Uniform nomenclature for bacterial plasmids: a proposal. *Bacteriol Rev*. 1976 Mar;40(1):168–89.
36. Poole P, Ramachandran V, Terpolilli J. Rhizobia: from saprophytes to endosymbionts. *Nat Rev Microbiol*. 2018;16(5):291-303.
37. Pinochet X, Arnaud F, Cleyetmarel JC. Competition for Nodule Occupancy of Introduced *Bradyrhizobium-japonicum* Strain Smgs1 in French Soils Already Containing *Bradyrhizobium-japonicum* Strain G49. *Can J Microbiol*. 1993;39(11):1022-8.
38. Ames P, Bergman K. Competitive Advantage Provided by Bacterial Motility in the Formation of Nodules by *Rhizobium-meliloti*. *J Bacteriol*. 1981;148(2):728-9.
39. Drogue B, Dore H, Borland S, Wisniewski-Dye F, Prigent-Combaret C. Which specificity in cooperation between phytostimulating rhizobacteria and plants? *Research in microbiology*. 2012;163(8):500-10.
40. Cieslinski G, Van Rees KCJ, Szmigielska AM, Krishnamurti GSR, Huang PM. Low-molecular-weight organic acids in rhizosphere soils of durum wheat and their effect on cadmium bioaccumulation. *Plant Soil*. 1998;203(1):109-17.
41. Owen AG, Jones DL. Competition for amino acids between wheat roots and rhizosphere microorganisms and the role of amino acids in plant N acquisition. *Soil Biol Biochem*. 2001;33(4-5):651-7.
42. Garcia JAL, Barbas C, Probanza A, Barrientos ML, Manero FJG. Low molecular weight organic acids and fatty acids in root exudates of two *Lupinus* cultivars at flowering and fruiting stages. *Phytochem Analysis*. 2001;12(5):305-11.
43. Ryan PR, Raman H, Gupta S, Horst WJ, Delhaize E. A second mechanism for aluminum resistance in wheat relies on the constitutive efflux of citrate from roots. *Plant Physiol*. 2009;149(1):340-51.
44. Liu MY, Lou HQ, Chen WW, Pineros MA, Xu JM, Fan W, et al. Two citrate transporters coordinately regulate citrate secretion from rice bean root tip under aluminum stress. *Plant Cell Environ*. 2018;41(4):809-22.
45. Canarini A, Kaiser C, Merchant A, Richter A, Wanek W. Root Exudation of Primary Metabolites: Mechanisms and Their Roles in Plant Responses to Environmental Stimuli. *Front Plant Sci*. 2019;10:157.
46. Glekas GD, Mulhern BJ, Kroc A, Duelfer KA, Lei V, Rao CV, et al. The *Bacillus subtilis* Chemoreceptor McpC Senses Multiple Ligands Using Two Discrete Mechanisms. *J Biol Chem*. 2012;287(47).

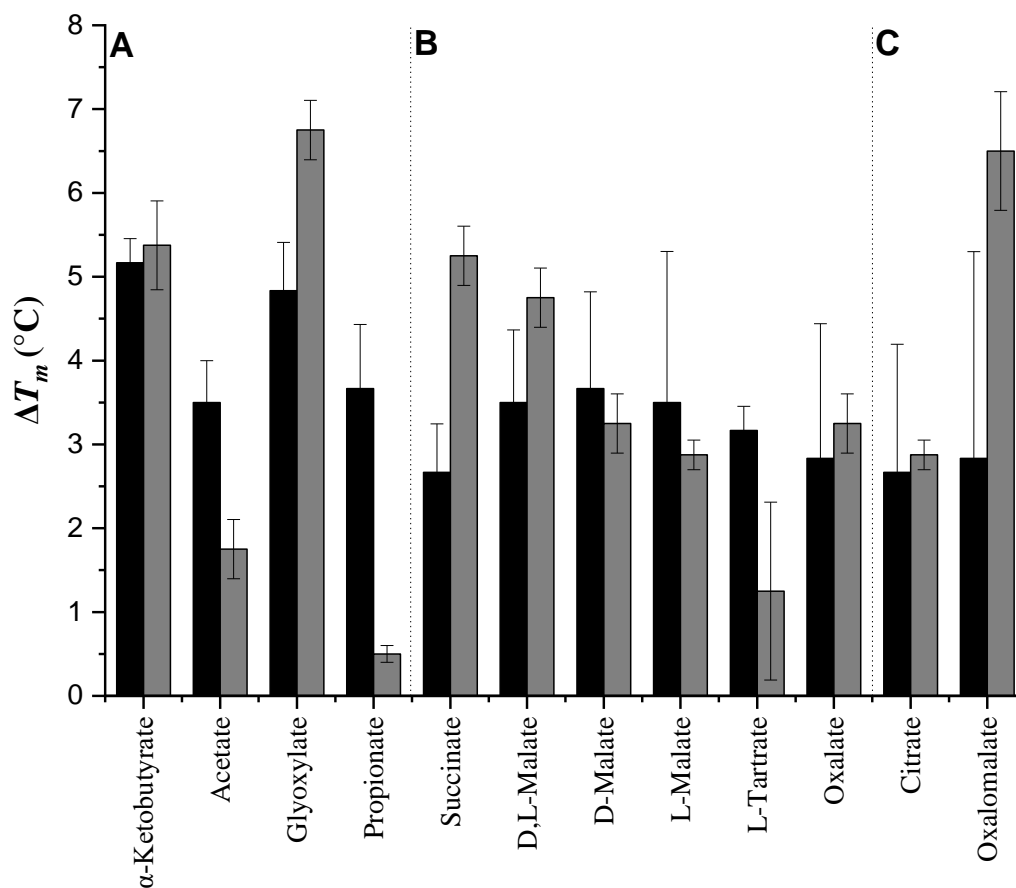
47. Schweinitzer T, Josenhans C. Bacterial energy taxis: a global strategy? *Arch Microbiol.* 2010;192(7):507-20.
48. Gherardi MJ, Rengel Z. The effect of manganese supply on exudation of carboxylates by roots of lucerne (*Medicago sativa*). *Plant Soil.* 2004;260(1-2):271-82.
49. Kamilova F, Kravchenko LV, Shaposhnikov AI, Azarova T, Makarova N, Lugtenberg B. Organic acids, sugars, and L-tryptophane in exudates of vegetables growing on stonewool and their effects on activities of rhizosphere bacteria. *Mol Plant Microbe Interact.* 2006;19(3):250-6.
50. Tao Q, Hou DD, Yang XE, Li TQ. Oxalate secretion from the root apex of *Sedum alfredii* contributes to hyperaccumulation of Cd. *Plant Soil.* 2016;398(1-2):139-52.
51. Kost T, Stopnisek N, Agnoli K, Eberl L, Weiskopf L. Oxalotrophy, a widespread trait of plant-associated Burkholderia species, is involved in successful root colonization of lupin and maize by Burkholderia phytofirmans. *Front Microbiol.* 2014;4.
52. Liu XL, Zhang KY, Liu YN, Xie ZH, Zhang CS. Oxalic Acid From *Sesbania rostrata* Seed Exudates Mediates the Chemotactic Response of *Azorhizobium caulinodans* ORS571 Using Multiple Strategies. *Front Microbiol.* 2019;10.
53. Dunn MF. Tricarboxylic acid cycle and anaplerotic enzymes in rhizobia. *Fems Microbiol Rev.* 1998;22(2):105-23.
54. Iyer B, Rajput MS, Jog R, Joshi E, Bharwad K, Rajkumar S. Organic acid mediated repression of sugar utilization in rhizobia. *Microbiol Res.* 2016;192:211-20.



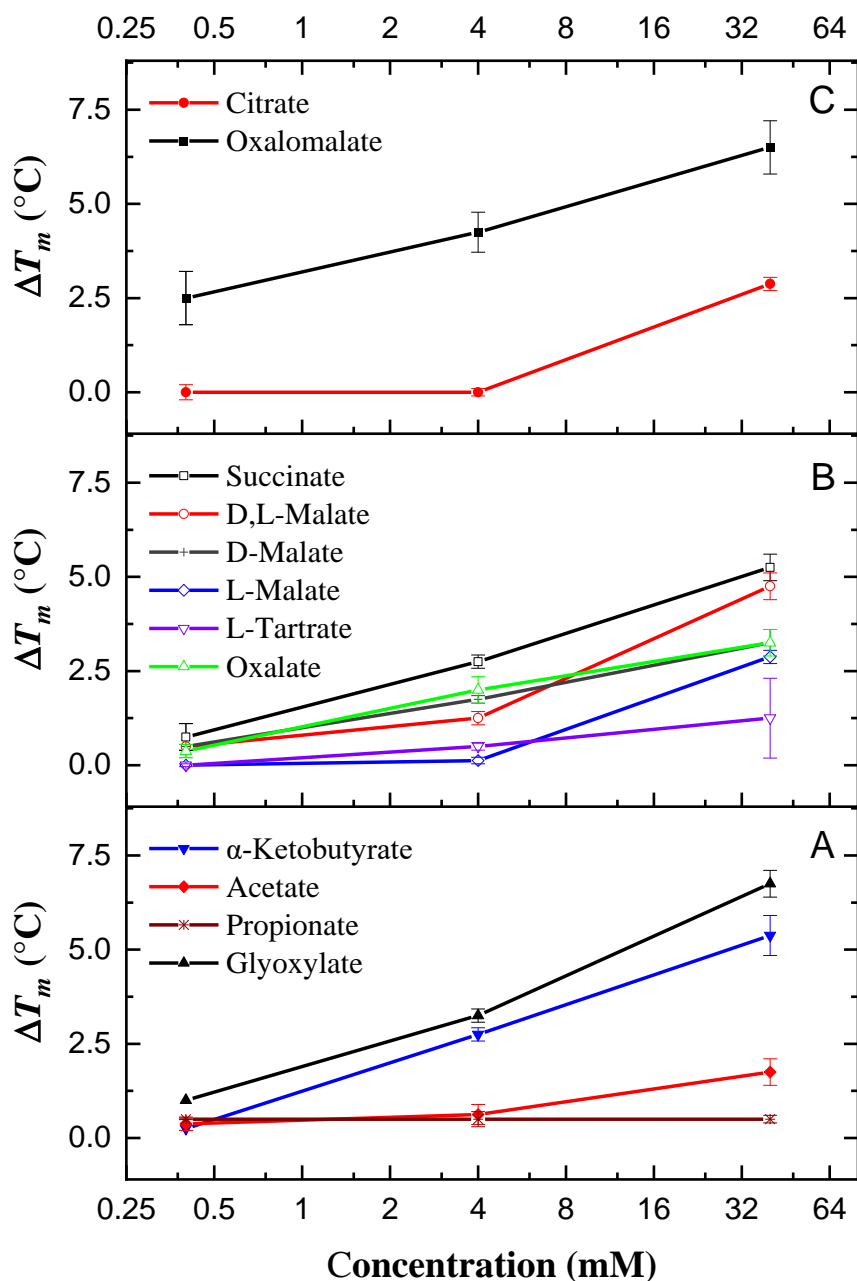
**Fig.4.1** Domain architecture of *Sinorhizobium meliloti* McpT, according to MIST and SMART databases. TM denotes the transmembrane regions; 4-HB is the 4-helix bundle sensor domain; HAMP is the conserved signal transduction domain; MH is a methyl-accepting helix; signaling is the MCP signaling domain that interacts with CheA and CheW.



**Fig 4.2.** High-throughput DSF screen of McpT<sup>PR</sup> with 190 potential ligands from Biolog plates PM1 (A) and PM2B (B). The  $\Delta T_m$  is the change in thermal stability of recombinant McpT<sup>PR</sup> in the presence of a compound.  $\Delta T_m$  values above the threshold (dotted line) of 2.5 °C indicate a possible ligand-protein interaction. The compounds identified as potential McpT-ligands are labeled. Values are the means and standard deviations from three biological replicates.

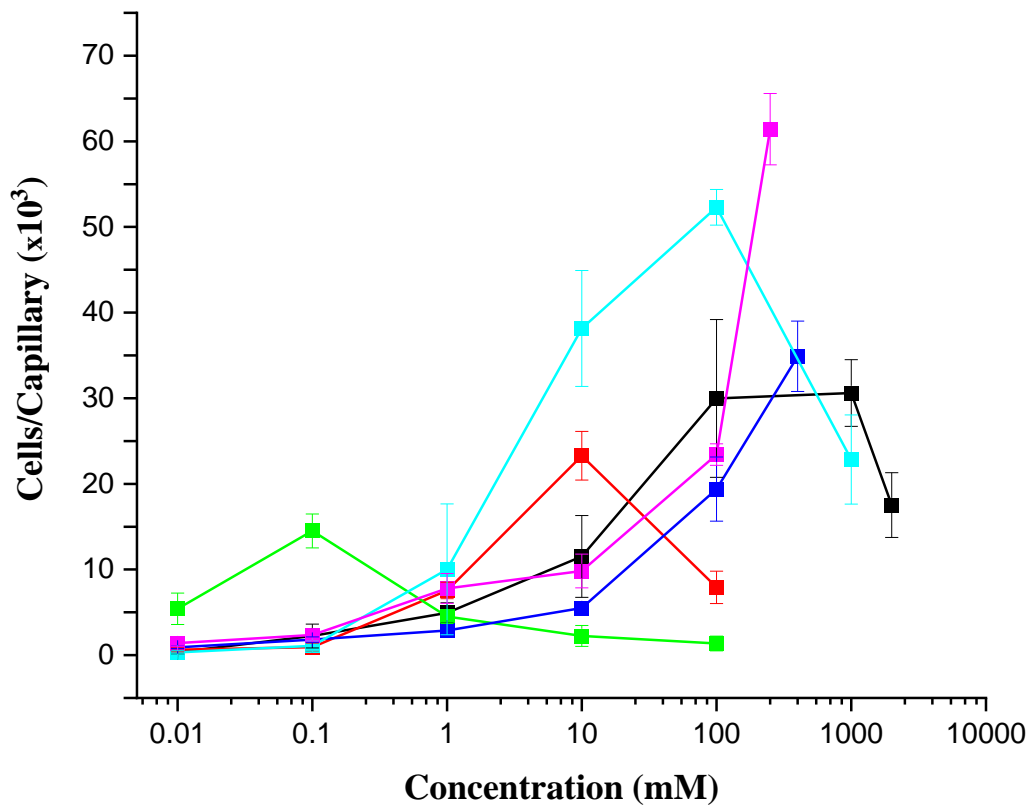


**Fig 4.3.** DSF screen of McpT<sup>PR</sup> with selected compounds from PM1 and PM2B Biolog plates. Data from Biolog plates (black) are compared to data with pure potential ligands at 40 mM (grey). (A) Mono-carboxylates, (B) Di-carboxylates, and (C) Tri-carboxylates. Only results with a  $\Delta T_m$  above 2.5 °C are presented. Values are the means and standard deviations from three biological replicates.

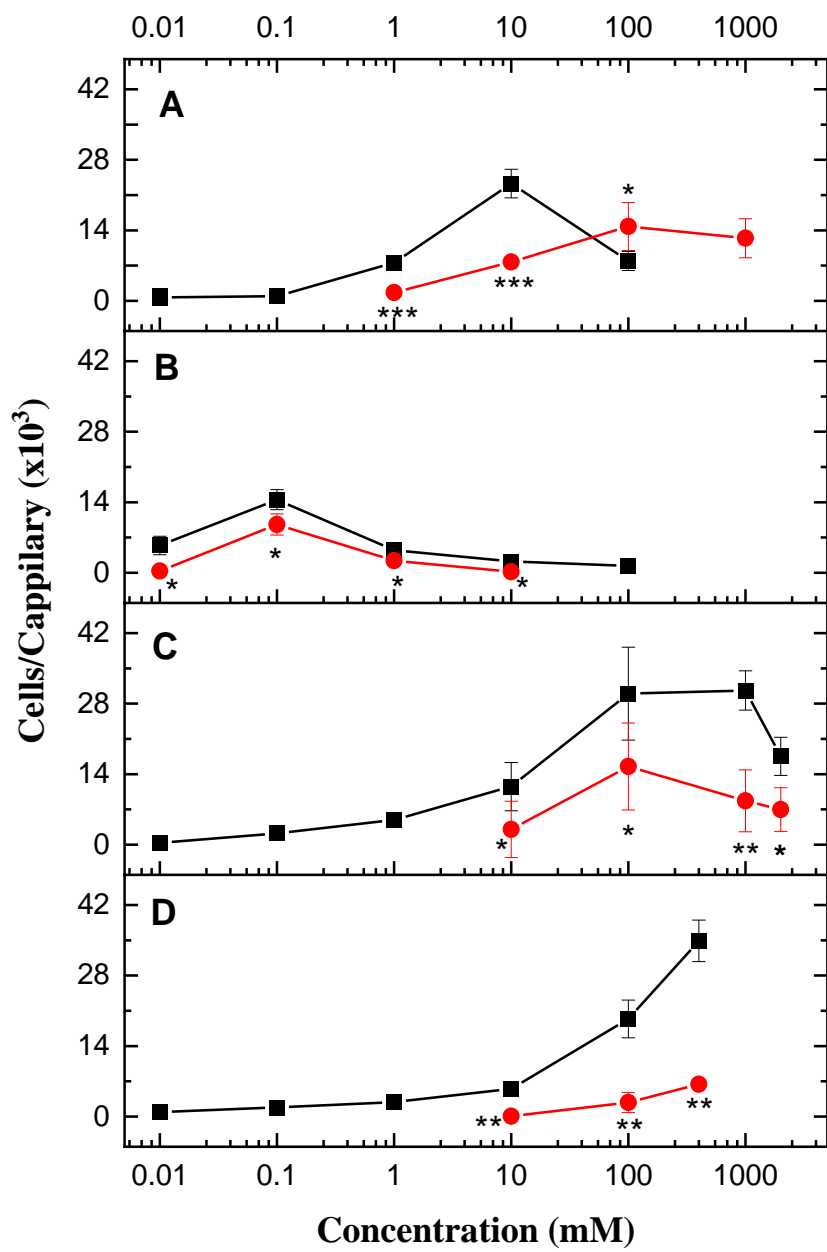


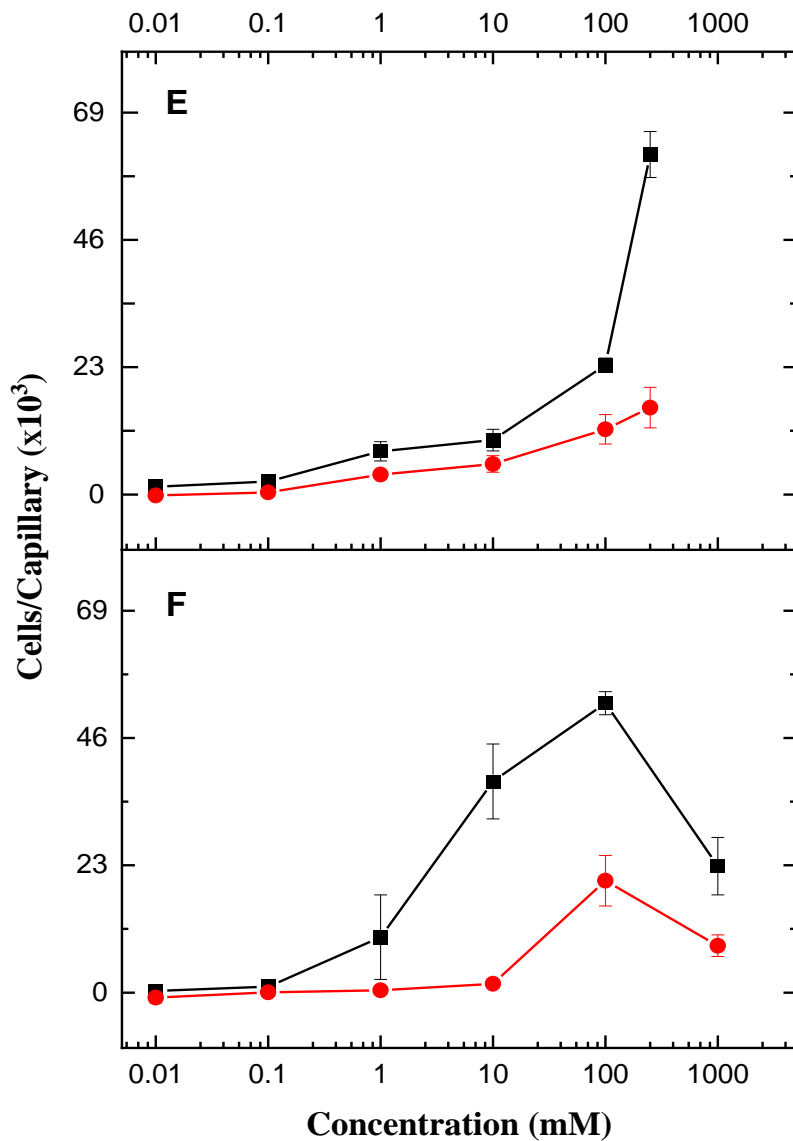
**Fig 4.4.** DSF screen of McpT<sup>PR</sup> with varying concentrations of potential ligands. (A) Mono-carboxylates, (B) Di-carboxylates, and (C) Tri-carboxylates. Specifically, compounds were tested at concentrations of 0.4, 4, and 40 mM. Data are the means and standard deviations from three biological replicates. If not visible, the error bars are contained within the symbol.



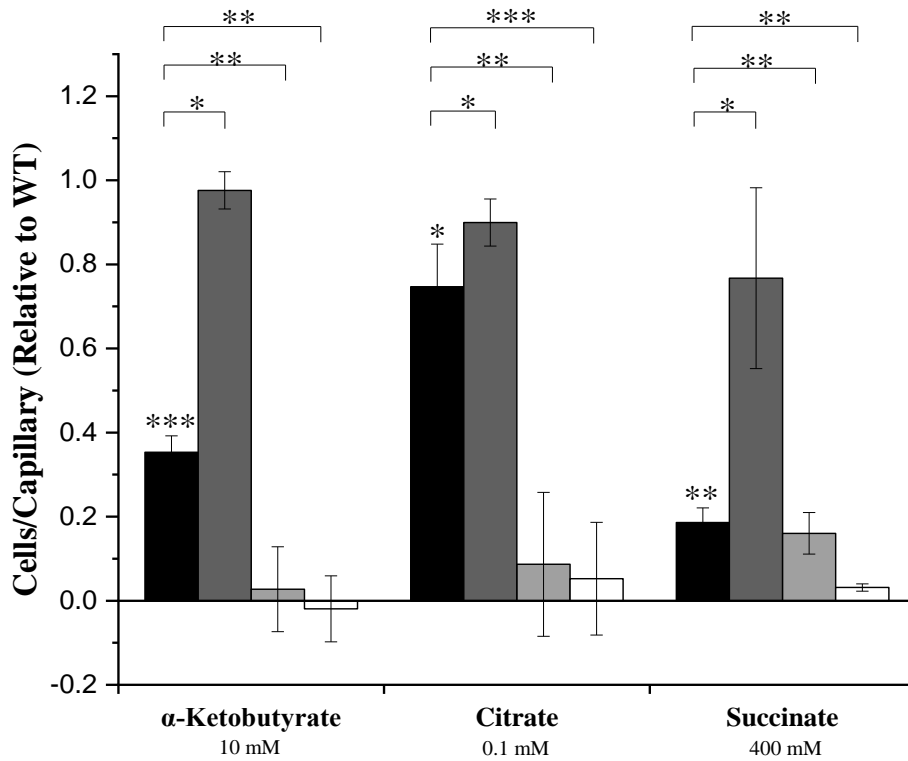


**Fig 4.5.** Chemotaxis response of *S. meliloti* wild type towards carboxylates in a quantitative capillary assay. Response curves to  $\alpha$ -ketobutyrate (red), citrate (green), glyoxylate (light blue) malate (black), oxalate (purple), and succinate (dark blue). The last data points for malate, oxalate, and succinate corresponds to 2,000, 250 and 400 mM, respectively. The data have been normalized to a negative control by subtracting the average number of cells that accumulated in control capillaries containing only RB medium. The data represent the mean and standard deviation of at least four independent biological replicates.

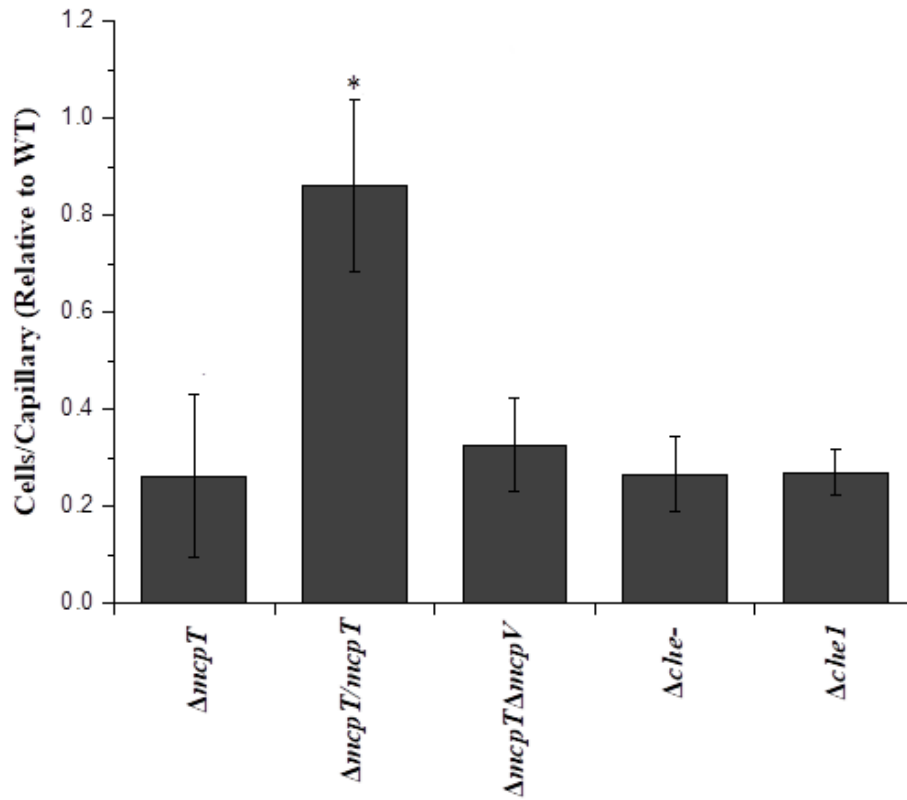




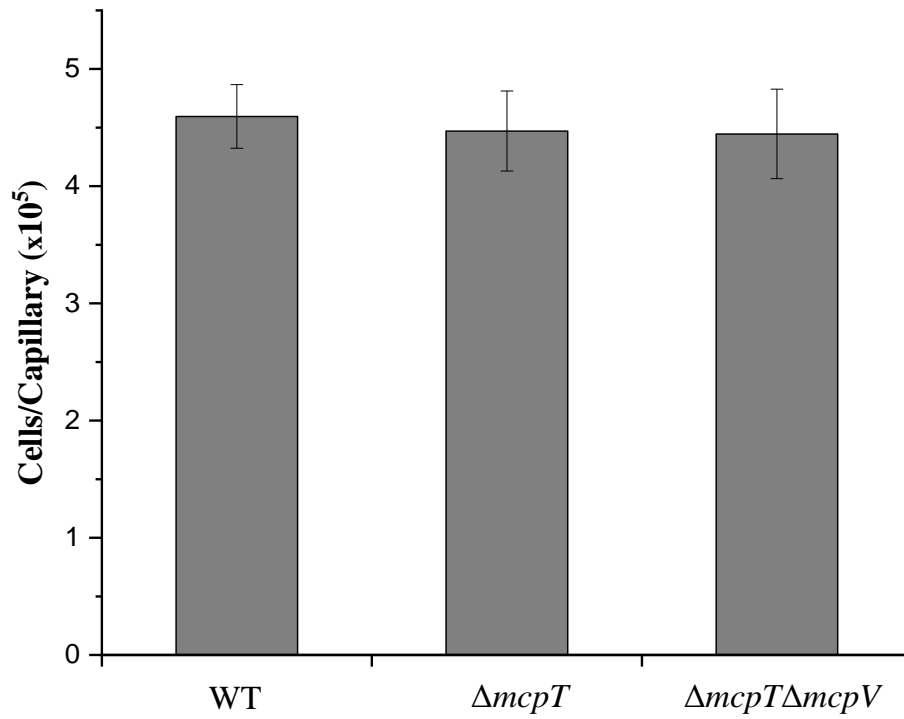
**Fig 4.6.** Chemotaxis responses of *S. meliloti* wild type (black) and a strain lacking *mcpT* (red) toward increasing concentrations of potential ligands. (A)  $\alpha$ -ketobutyrate, (B) citrate, (C) malate, (D) succinate, (E) oxalate, and (F) glyoxylate. Chemotaxis data of the wild-type response are taken from Fig. 4.3. Values are the means and standard deviations of at least four biological replicates. If not visible, the error bars are contained within the symbol. Note the difference in scale between the two panels. Values are the means and standard deviations from three biological replicates. Asterisks denotes *P*-values determined by a Student *t* test: \*,  $p < 0.05$ ; \*\*,  $p < 0.01$ ; \*\*\*,  $p < 0.0001$ .



**Fig 4.7.** Chemotactic responses of various *S. meliloti* deletion strains to selected carboxylates relative to wild type. Black bars,  $\Delta mcpT$  (RU11/838); dark grey bars,  $\Delta mcpT/mcpT$  (RU111/838 with pBS1055); light grey bars,  $\Delta mcpT \Delta mcpV$  strain (BS275); white bars, *che<sup>-</sup>* strain (RU13/149). The mean of cell numbers per capillary for each strain was normalized to that of wild-type. Values are the means and standard deviations of at least four biological replicates. Asterisks denotes *P*-values determined by a Student *t*-test: \*, *p* <0.05; \*\*, *p* <0.01; \*\*\*, *p* <0.0001. *P*-values above  $\Delta mcpT$  represents the significant differences compared with the wild type.

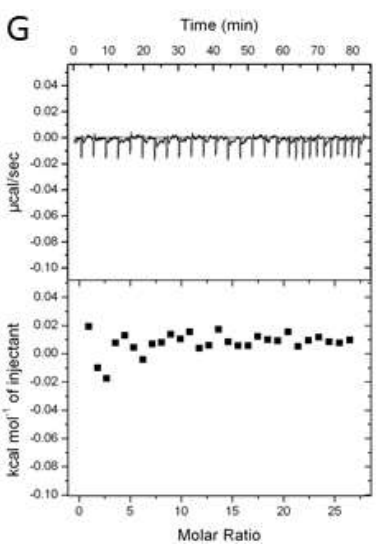
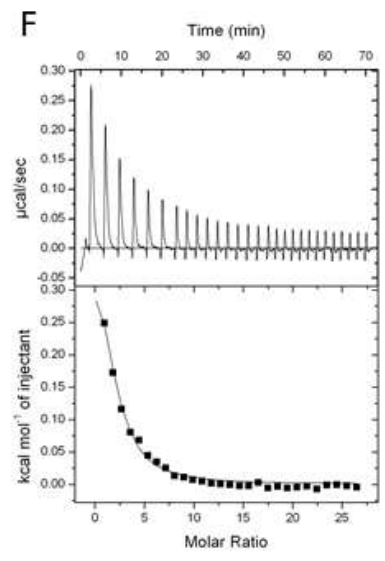
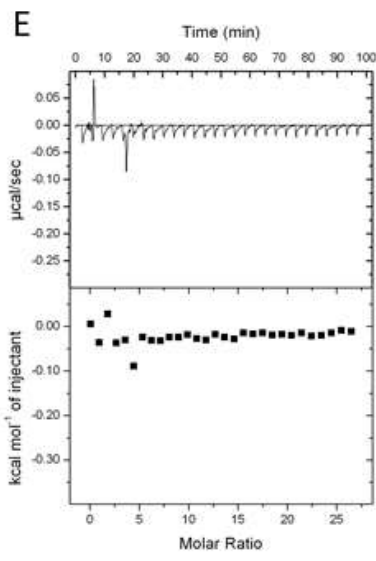
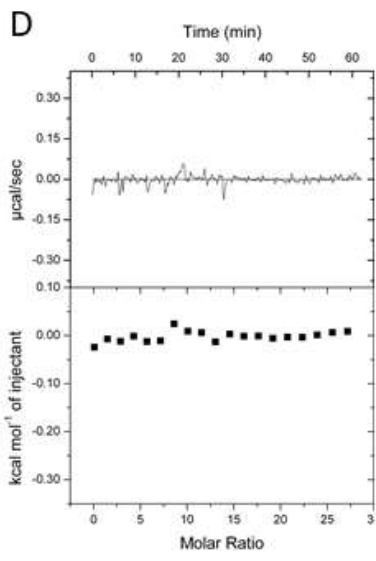
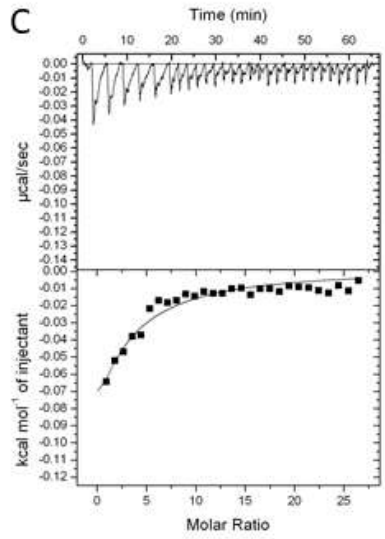
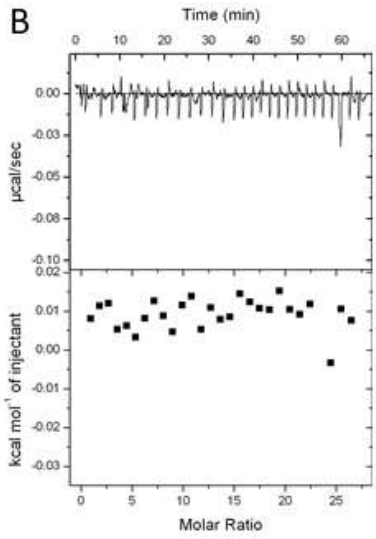
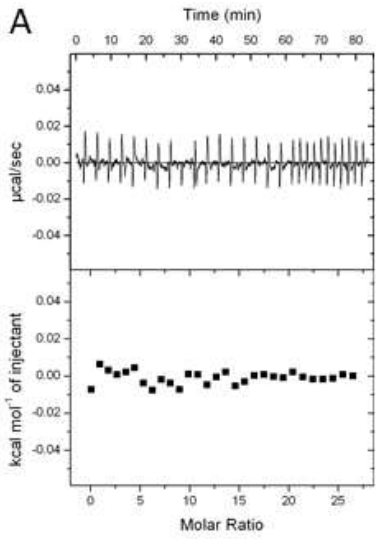


**Fig 4.8.** Chemotactic responses of various *S. meliloti* deletion strains to the peak concentration of malate (1000 mM). The mean of cell numbers per capillary for each strain was normalized to that of wild-type. Values are the means and standard deviations of at least four biological replicates. Asterisks denotes *P*-values determined by a Student *t*-test: \*, *p* < 0.05.



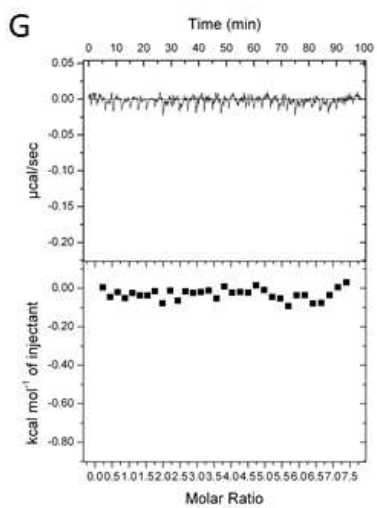
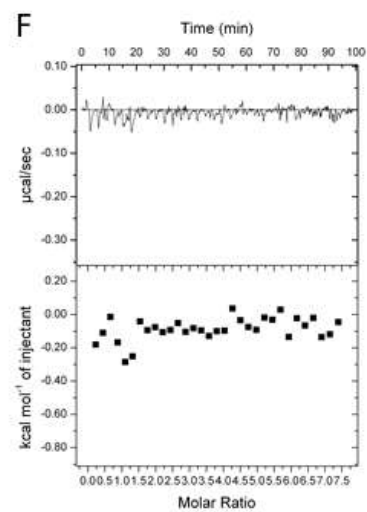
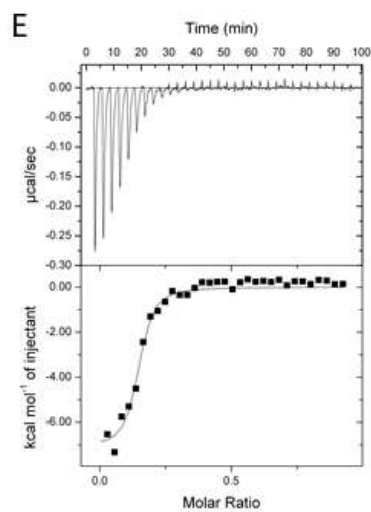
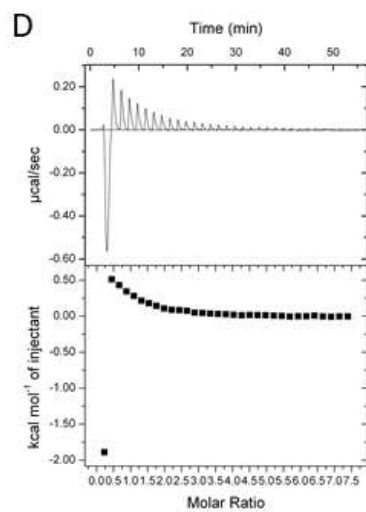
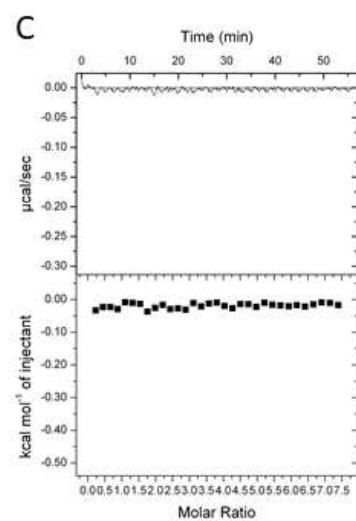
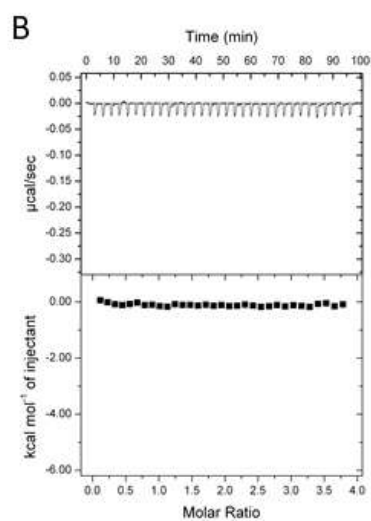
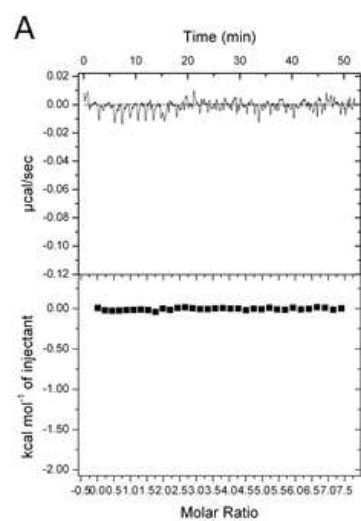
**Fig**

**4.9.** Chemotaxis responses of *S. meliloti* wild type,  $\Delta mcpT$ , and  $\Delta mcpT\Delta mcpV$  to 10 mM proline. Values are the means and standard deviations from three biological replicates.

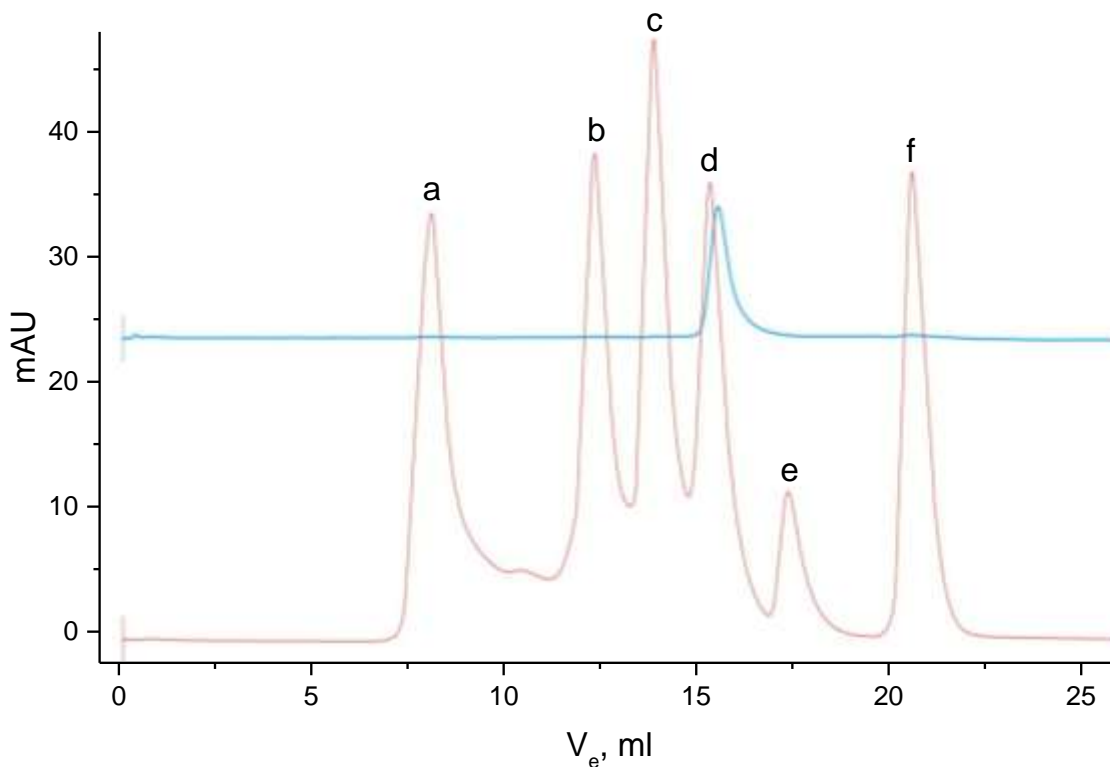


**Fig. 4.12** Microcalorimetry titrations of 40  $\mu\text{M}$  McpT<sup>PR</sup> (except for (D), which used 70  $\mu\text{M}$ ) in 0.1 M Tricine, 0.2 M NaCl, 0.2 M KCl, 15 % glycerol, pH 8.0 against putative ligands at 5 mM. Each top panel shows the thermogram, each bottom panel the integrated heat signals. The curves of best fit were created using the “one binding site” function in the MicroCal version of Origin 7. (A) citrate ; (B) glycolate ; (C) glyoxylate ; (D) malate ; (E) oxaloacetate ; (F) oxalate; (G) succinate.

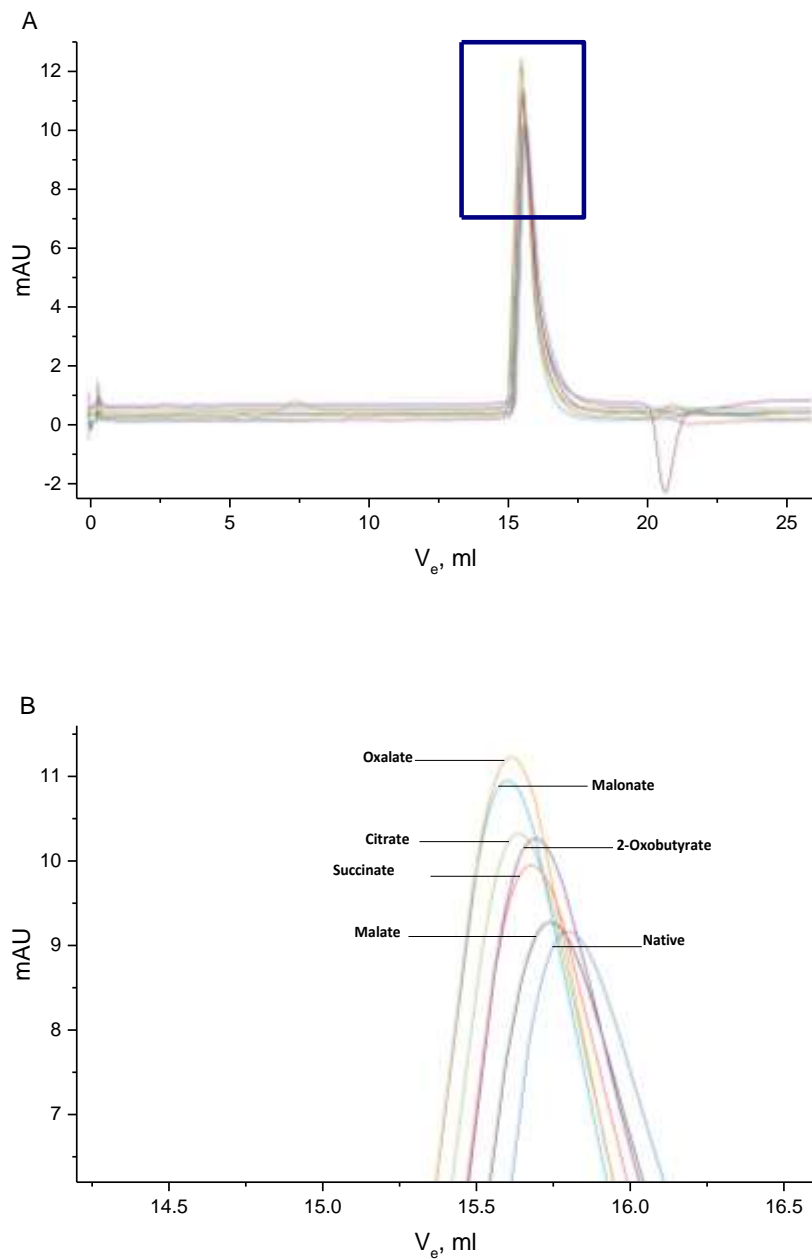




**Fig. 4.11** Microcalorimetry titrations of 45  $\mu\text{M}$  McpT<sup>PR</sup> in 0.1 M PIPES buffer, pH 7.0 against putative ligands at 2 mM (except for (E), which used 250  $\mu\text{M}$ ). Each top panel shows the thermogram, each bottom panel the integrated heat signals. The curves of best fit were created using the “one binding site” function in the MicroCal version of Origin 7. (A) citrate ; (B) glyoxylate ; (C) malate ; (D) malonate ; (E) 250  $\mu\text{M}$  oxalate ; (F) succinate ; (G) alpha-ketobutyrate.



**Fig. 4.12** Analytical size exclusion chromatography of McpT<sup>PR</sup> (blue) compared to molecular weight standards (orange) in 0.1 M PIPES, 0.1 M KCl [pH 7.0]. Samples were applied to the column in 200  $\mu$ l injections. a. Dextran blue, 2,000 kDa; b. Aldolase, 158 kDa; c. Conalbumin, 75 kDa; d. Ovalbumin, 44 kDa; e. Cytochrome C, 12.4 kDa ; f. Vitamin B12, 1.4 kDa.



**Fig. 4.13** Analytical size exclusion chromatography of McpT<sup>PR</sup> in the presence of putative ligands. Samples were applied in 200  $\mu$ l injections of titrated protein. McpT<sup>PR</sup> was separated in 0.1 M PIPES, 0.1 M KCl [pH 7.0], amended with 5 mM of each putative ligand. A. Full range chromatogram with McpT<sup>PR</sup> eluting between 15 and 16 ml. The box marks the zoomed section shown in B.

**Chapter 5 - Structure of the sensory domain of McpX from *Sinorhizobium meliloti*, the first known bacterial chemotactic sensor for quaternary ammonium compounds**

Manisha Shrestha<sup>1</sup>, Karl K. Compton<sup>1</sup>, Jordan M. Mancl<sup>1</sup>, Benjamin A. Webb<sup>1</sup>, Anne M. Brown<sup>2</sup>, Birgit E. Scharf<sup>1</sup> and Florian D. Schubot<sup>1</sup>

<sup>1</sup>Department of Biological Sciences, Virginia Tech, Derring Hall, Blacksburg, VA 24061, U.S.A.; <sup>2</sup>Department of Biochemistry, Virginia Tech, Engel Hall, Blacksburg, VA 24061, U.S.A.

Correspondence: Birgit E. Scharf (bscharf@vt.edu) or Florian D. Schubot (fschubot@vt.edu)

**Biochemical Journal. *Biochem J* (2018) 475 (24): 3949–3962. Published online 14<sup>th</sup> of December, 2018. <https://doi.org/10.1042/BCJ20180769>**

**Attribution: M.S. performed the structural studies and analysis, K.K.C. isolated proteins and performed the ITC studies, J.M.M. obtained the initial crystals for McpX, assisted with the model building and manuscript, B.A.W. developed protocols for McpX purification and characterization, A.M.B. performed the molecular docking studies, B.E.S. co-wrote the manuscript, F.D.S. supervised all structural and biochemical analysis, and co-wrote the manuscript.**

## ABSTRACT

The  $\alpha$ -proteobacterium *Sinorhizobium meliloti* can live freely in the soil or engage in a symbiosis with its legume host. *S. meliloti* facilitates nitrogen fixation in root nodules, thus providing pivotal, utilizable nitrogen to the host. The organism has eight chemoreceptors, namely McpT to McpZ and IcpA that facilitate chemotaxis. McpX is the first known bacterial sensor of quaternary ammonium compounds (QACs) such as choline and betaines. Because QACs are exuded at chemotaxis-relevant concentrations by germinating alfalfa seeds, McpX has been proposed to contribute to host-specific chemo-taxis. We have determined the crystal structure of the McpX periplasmic region (McpXPR) in complex with the proline betaine at 2.7 Å resolution. In the crystal, the protein forms a symmetric dimer with one proline betaine molecule bound to each monomer of McpXPR within membrane-distal CACHE module. The ligand is bound through cation- $\pi$  interactions with four aromatic amino acid residues. Mutational analysis in conjunction with binding studies revealed that a conserved aspartate residue is pivotal for ligand binding. We discovered that, in a striking example of convergent evolution, the ligand-binding site of McpXPR resembles that of a group of structurally unrelated betaine-binding proteins including ProX and OpuAC. Through this comparison and docking studies, we rationalized the specificity of McpXPR for this specific group of ligands. Collectively, our structural, biochemical, and molecular docking data have revealed the molecular determinants in McpX that are crucial for its rare ligand specificity for QACs.

## INTRODUCTION

Bacteria utilize a wide range of naturally occurring organic compounds as sources for carbon, nitrogen, and for the fundamental building blocks required for the synthesis of proteins, lipids, and nucleic acids. To identify and seek out diverse nutrient sources and eukaryotic hosts, many bacteria have evolved a mechanism to link their swimming motility with a complex chemosensory system [1– 3]. This process, called chemotaxis, entails the coupling of the chemotactic sensory machinery with the rotation of the flagella and thus swimming motility. Chemotactic behavior is best understood in *Escherichia coli*, which uses several peritrichous flagella to form a bundle for propulsion [4]. The canonical sensing mechanism for environmental signals involves signal recognition by chemoreceptors called methyl-accepting chemotaxis proteins (MCPs), which typically consist of two transmembrane domains, a variable periplasmic ligand-binding domain, and a conserved cytoplasmic signaling domain [5]. MCPs form a ternary complex with the adaptor protein CheW that mediates binding the MCPs to the histidine autokinase CheA [6]. In the absence of a chemoattractant or in the presence of a repellent, CheA autophosphorylates and then transfers the phosphate group to its cognate response regulator CheY. Phosphorylated CheY binds to the flagellar motor to cause a disruption of the flagellar bundle resulting in bacterial cell tumble and thus random reorientation [7,8]. Binding of a chemoattractant to its cognate MCP induces a conformational change in the sensory domain to produce a piston-type motion accompanied by a twisting motion across the cellular membrane, which alters the conformation of the kinase control region. This structural change reduces the activity of CheA, thus slowing the rate of CheY phosphorylation and thereby reducing the tumble frequency of the cell. The overall outcome is a biased random walk of the bacterial cell toward the source of the detected attractant or away from a repellent [5].

MCPs are obligate dimers, which form trimers of dimers and further organize into large-scale arrays at the cell pole that are composed of MCPs with differing ligand specificities [9]. These arrays serve to amplify the signal originating from individual ligand bound MCPs to increase the sensitivity of the system [10]. The associated conformational change cascades across the array to induce similar structural changes in receptors without a bound ligand [11]. The chemotactic response is further modulated through the reversible methylation of conserved sites in the cytosolic region of MCPs. A constitutively active methyl-transferase adds methyl groups, while a

methylesterase removes methyl groups after it is activated through phosphorylation by CheA. Methylation creates a memory of recent ligand-binding events and provides an adaptive mechanism to prevent saturation of the chemotaxis system as the bacterium swims up a nutrient gradient. Attractant binding reduces CheA activity but also stimulates the gradual methylation of the MCP, which in turn increases CheA activity leading to a gradual recovery of the tumble frequency [12]. This adaptive mechanism thereby reduces the sensitivity of the MCP array and ensures that only a further increased chemoattractant concentration will produce a chemotactic response to guide the bacterial cell further up the gradient.

Chemotaxis enables the soil-dwelling plant symbiont *Sinorhizobium meliloti* to actively seek the suitable site for host root infection [13–15]. It possesses eight distinct chemoreceptors sequentially named McpT to McpZ and IcpA, which lacks the conserved methylation sites, to sense environmental cues [16]. IcpA and McpY do not have transmembrane regions but still associate with the chemoreceptor cluster at the cell poles [17]. It is hypothesized that these two chemoreceptors sense compounds that have been internalized or could be used as energy level sensors by somehow sensing the energy flux in the cell [15]. Signal sensing for the six transmembrane MCPs is based on chemoreception in the periplasm. The ligand-binding specificity of two of the eight *S. meliloti* chemoreceptors involved in chemotaxis has been determined. McpU is a general amino acid receptor, sensing all non-acidic proteogenic amino acids, as well as several non-proteogenic amino acids [18–20]. McpX senses quaternary ammonium compounds (QACs) such as glycine betaine, proline betaine, trigonelline, and choline [21]. *S. meliloti* utilizes proline betaine as a nutrient source but various microorganisms accumulate these compounds inside the cell to counter stress caused by high osmolarity and extremes in growth temperature [22]. Amino acids and QACs are commonly exuded by plant seeds and roots including those of alfalfa, the plant symbiont of *S. meliloti* [19,21]. The periplasmic regions of MCPs may be classified according to the predicted folds of their domains, and in some instances, a particular domain fold may be predictive of the MCPs specificity. The periplasmic regions of McpU and McpX were predicted to each contain a dual CACHE domain (dCACHE\_1), an arrangement frequently encountered in amino acid sensors [23]. Mutational analyses and molecular modeling have shown that amino acid ligands bind to the amino-proximal CACHE module in McpU [20]. Profiling of the ligand specificity of McpX revealed that, in addition to QACs, it only binds one amino acid with



significant affinity, namely proline [21,24–26]. Therefore, the ligand specificity of McpX may play a role in promoting symbiosis.

In the present study, we report the crystal structure of the McpX ligand-binding periplasmic region (McpXPR) in complex with the betaine proline betaine. The structure revealed striking similarities between the ligand-binding pocket of McpX and that found in a structurally unrelated group of QAC-binding proteins. Our structural analysis was used to explain how McpX reconciles having a strong selectivity for QACs with maintaining a broad affinity for compounds as chemically distinct as choline and proline betaine. Collectively, this work uncovered a remarkable example of convergent evolution and offers intriguing insights into nature's design principles for achieving specificity through subtle modifications in binding pocket properties. Ultimately, understanding these principles may enable us rationally engineer novel specificities to gain control of bacterial chemotaxis.

## MATERIALS AND METHODS

### Chemicals

Proline betaine (L-proline betaine) was purchased from Extrasynthese (Toulouse, France) and choline from Sigma-Fluka (St. Louis, MO, U.S.A.).

### Expression and purification of the periplasmic region of McpX (McpXPR) and crystallization

The *mcpX* 100–919 bp fragment was amplified with Phusion DNA polymerase (NEBiolabs) using chromosomal DNA as template and cloned into Qiagen expression vector pQE30 using BamHI and HindIII sites to produce the expression plasmid pBS455, wherein an N-terminal His6-tag is fused in frame with the codons for amino acid residues 34–306 of the *mcpX* gene. The clone was verified using DNA sequencing with pQE30-specific primers. The recombinant ligand-binding, a periplasmic region of McpX (McpXPR, McpX34–306) was overproduced from plasmid pBS455 in *E. coli* M15/pREP4 cell. Four liters of cell culture was grown to an OD<sub>600</sub> of 0.7 at 37°C in LB containing 100 µg ml<sup>-1</sup> ampicillin and 25 µg ml<sup>-1</sup> kanamycin. Gene expression was induced by the addition of 0.6 mM isopropyl-β-D-thiogalactopyranoside for 4 h at 25°C. Cells were harvested via centrifugation and stored at -30°C. Cells were suspended in 70 ml column buffer [500 mM NaCl, 25 mM imidazole, 20 mM NaPO<sub>4</sub>, pH 7.4, 2 mM tri(2-carboxyethyl)phosphine (TCEP), 1 mM phenylmethylsulfonyl fluoride (PMSF)] with 1 µg/ml of DNase and lysed by three passages

through a French pressure cell at 20 000 psi (SLM Aminco, Silver Spring, MD, U.S.A.). The soluble fraction was loaded onto three stacked 5 ml NTA columns (GE Healthcare Life Sciences) charged with Ni<sup>2+</sup>. Protein was eluted from the column in a linear gradient of elution buffer (500 mM NaCl, 350 mM imidazole, 20 mM NaPO<sub>4</sub>, pH 7.0, 2 mM TCEP, 1 mM PMSF). Protein-containing fractions were pooled and loaded onto a HiPrep 26/60 Sephacryl S-300 HR (GE Healthcare) gel filtration column. Prior, this column had been equilibrated in column buffer containing 100 mM NaCl, 50 mM HEPES, pH 7.0. Peak fractions were analyzed via SDS-PAGE. Fractions containing >95% purified McpXPR were pooled and concentrated via ultrafiltration with 10-kDa cutoff regenerated cellulose membranes in a 50 ml Amicon filter unit (Millipore, Bedford, MA, U.S.A.), and stored at 4°C.

The change of Asp208 to Asn was introduced into mcpX using overlap extension PCR followed by the same cloning steps that had been used for the generation of pBS455 except that the parental plasmid was pQE60, creating pBS517. Subsequent expression and purification steps of the Asp208Asn McpXPR variant (McpXD208N-PR) mirrored those used for the purification of the original McpXPR protein. McpXD208N-PR behaved identically to the original protein during sample purification and concentration. In a preliminary differential scanning fluorimetry experiment, the protein did not show binding to proline betaine but produced a sharp melting transition curve characteristic of a folded and homogeneous sample (not shown).

Initial crystals of McpXPR in the presence of proline betaine were obtained through high-throughput screening of commercial crystallization conditions using a sitting drop format and a Honeybee961 crystallization robot. Optimized crystals were obtained in a hanging drop set-up by combining 3 ml protein solutions containing 190 mM McpXPR, 9 mM proline betaine, 100 mM NaCl, and 50 mM HEPES, pH 7.0 with 1 ml of a screen-ing condition containing 0.08 M sodium acetate, 1.6 M ammonium sulfate, and 20% glycerol, pH 4.6. The droplet was allowed to equilibrate at room temperature against a reservoir containing 64 mM sodium acetate, 1.28 M ammonium sulfate, and 16% glycerol, pH 4.6. Crystals appeared within a day. Many crystals displaying a wheel-like shape proved to be twinned but others, displaying a topology more reflective of their tetrahedral lattice system, yielded untwinned diffraction data to ~2.7 Å. X-ray diffraction data collection, structure determination, and refinement X-ray data at 2.7 Å were collected at SER-CAT 22-ID at the Advanced Photon Source in Argonne, Il using a

MARMOSAIC 300 CCD detector. The X-ray diffraction data were processed using the HKL2000 program package. Data were analyzed with Xtriage from the PHENIX program suite [27] to confirm that the crystals were not twinned. Initial phases were obtained with molecular replacement using the structure of the extracellular domain of the putative histidine kinase mmHK1S-Z3 as a search model (PDB code: 3lib). PHASER from within the PHENIX program suite was used for the search and to calculate the initial density map. Model building was performed using COOT, and the PHENIX program suite was used for model refinement [27]. During the refinement, the diffraction data were cut off at 2.7 Å using the correlation of the observed data set with the refined model, CC1/2, as defined by Karplus and Diederichs as selection criterion [28]. Data collection and refinement statistics are summarized in Table 1. The refined model was deposited in the protein data bank under the accession code 6D8V.

#### Isothermal titration calorimetry

McpXPR or McpXD208N-PR in 100 mM NaCl and 50 mM HEPES, pH 7.0 was used at 20 µM for the ITC (isothermal titration calorimetry) measurements with choline and proline betaine. Both ligands were dissolved in dialysis buffer. Measurements were performed with a VP-ITC Microcalorimeter (MicroCal, Northampton, MA, U.S.A.) at 15°C. McpXPR or McpXD208N-PR was placed in the sample cell. Baselines were produced by titrating the ligands into dialysis buffer void of protein. These baselines were subtracted from each protein titration. Data analysis was carried out with the MicroCal version of Origin 7.0 software using the ‘one binding sites’ model. (Origin Laboratory, Northampton, MA, U.S.A.).

#### Molecular docking

Molecular docking of the bound crystal structure ligand (proline betaine) and five additional ligands (betonicine, choline, glycine betaine, trigonelline, and proline) was performed to validate the protocol and explore binding pocket specificity of McpXPR. The crystal structure of McpXPR was used as the receptor, with water molecules removed before receptor preparation in AutoDock Tools (ADT) [29]. Autodock Vina [30] was used to perform the docking and pose prediction, The same grid box co-ordinates (23 Å × 20 Å × 20 Å) and center (0.437, 0.303, 0.167) with a 1.000 Å grid spacing was used to dock all ligands and was based on the position of proline betaine in the solved structure. Nine poses for each ligand docked to McpXPR were generated. Our box size, center, and protocol were validated with the re-docking of proline betaine in the

solved structure, producing a low-energy pose of  $-6.2$  kcal/mol and a root-mean-square deviation (RMSD) of  $1.102$  Å between solved and docked proline betaine position. Notably, for all additional ligands tested, the top six energetically favorable poses had the quaternary amine group clustered at near the position as that of proline betaine. The lowest energy pose for each ligand docked was selected for further analysis to determine key interactions and residues for functionality. Distances were assessed to predict hydrogen bonding (less than  $3.5$  Å), electrostatic ( $2.5$ – $3.5$  Å), hydrophobic ( $3.4$ – $3.9$  Å), and distal, weaker interactions (greater than  $4.0$  Å) to determine the rationale for ligand specificity.

## RESULTS

The periplasmic McpX sensory region assumes a canonical dual CACHE domain fold

The overall structure of McpXPR assumes a dual Calcium channels and Chemotaxis receptor (dCACHE\_1) fold, wherein a membrane-proximal and a membrane-distal module are folded against the N-terminal and C-terminal halves of a long stalk helix  $\alpha 1$ , respectively (Figure 5.1A). The membrane-distal module (residues 65–210) contains a six-stranded, antiparallel  $\beta$ -sheet, which is flanked on one side by helix  $\alpha 2$  and the C-terminal half of the stalk helix, and on the other side by helix  $\alpha 3$ . On the other side of the  $\beta$ -sheet, helix  $\alpha 3$  packs against  $\beta$ -strands  $\beta 4$ ,  $\beta 5$ , and  $\beta 6$  to create part of the proline betaine-binding pocket. The membrane-proximal module (residues 38–63, 212–306) is formed by a five-stranded antiparallel  $\beta$ -sheet, which is flanked on one side by helix  $\alpha 4$  and the N-terminal half of helix  $\alpha 1$ , and by helix  $\alpha 5$  on the other side. As has been observed in structurally related proteins, the membrane-distal and membrane-proximal modules are packed tightly against each other, suggesting an arrangement that is likely important for communicating signal binding in the distal module across the inner membrane. Although the asymmetric unit of the crystal is formed by a single MCPXPR molecule, application of a crystallographic symmetry axis creates a dimer (Figure 5.1B) that closely resembles the dimers observed for structurally homologous MCP ligand-binding domains [31–33]. The extensive dimer interface is primarily composed of a four-helix bundle formed by helix  $\alpha 2$  and the C-terminal half of helix  $\alpha 1$ , and the equivalent regions of a symmetry-related molecule to create the dimerization interface burying a total surface area of  $6836$  Å<sup>2</sup> between the two molecules. A striking feature of the dimer is the approximate 25-degree kink in the stalk helix between residues 25 and 26 (Figure 5.1C). This kink leads to an increase in the distance between the symmetry-related stalk helices

from less than 5 Å between the  $\alpha$ -carbons of the symmetry-related Glu-65 and Glu-650 residues just above the kink to more than 22 Å between the  $\alpha$ -carbons of Arg-38 and Arg-380 at the amino-terminal ends of the stalk helices. The kinking has been pro-posed previously to be the result of ligand binding in the membrane-distal CACHE region as a means to facilitate signal transduction upon ligand binding [31].

The ligand-binding pocket is located in the distal CACHE module

As the refinement progressed, we inspected the structure of the membrane-distal module for additional electron density in a groove that bound the small-molecule ligands in other dCACHE\_1 containing receptors [31–35]. The membrane-proximal CACHE module has generally been proposed to facilitate signal transmission from the distal CACHE module to the membrane-bound regions of the receptor [1]. However, because there has been at least one instance where a small-molecule ligand bound to the membrane-proximal CACHE region [36], we also examined its putative binding pocket in McpXPR. Both the weighted  $2F_o - F_c$  and the  $F_o - F_c$  electron density maps produced strong residual density in the distal CACHE region inside a groove that is formed by beta strands  $\beta_1$ ,  $\beta_4$ ,  $\beta_5$ , the loop region connecting strand  $\beta_3$  and helix  $\alpha_3$ , and a ‘tongue-like loop’ encompassing residues 175–186 that connects strands  $\beta_5$  and  $\beta_6$  (Figures 5.2 and 5.4A). Modeling proline betaine into this density produced an excellent initial fit as judged by visual inspection. Subsequently, PHENIX was used to optimize the positioning and refine the ring conformation. In the resulting complex, the proline betaine molecule occupies the equivalent space as the serine and alanine in structures of Mlp37 from *Vibrio cholerae* [33] (PDB code: 5avf and 3c8c, respectively, Supplementary Figure 5.6 A and B), as the isoleucine ligand in the *Campylobacter jejuni* chemoreceptor Tlp3 [32] (PDB code: 4xmr, Supplementary Figure 5.6C), as Bis-Tris in the putative histidine kinase mmhk1s-z2 from *Methanosarcina frisia* [31] (PDB code: 3LIA, Supplementary Figure 5.7D), as putrescine in McpU from *Pseudomonas putida* (PDB code: 6F9G) [37] and as the arginine-ligand in the *P. aeruginosa* receptor protein PctB (PDB code: 5LT9, unpublished; Figure 5.6E). One of the primary objectives of this work was to gain an understanding of the molecular basis for the unusual specificity of McpX for QACs. The experimentally determined model revealed that proline betaine is coordinated on four sides by aromatic residues, namely Trp109, Tyr139, Phe153, and Trp161. The planes of the aromatic rings are oriented toward a positively charged amine group of the ligand creating four sets of cation–

$\pi$ interactions. The carboxylate group of the ligand is stabilized by hydrogen-bonding interactions with the backbone amide groups of Ala179 and Glu180. A similar arrangement was only observed in the structure of the Bis-Tris complex of Mmhk1s-z2, where the ligand is also caged-in on four sides by aromatic residues. Three of the four tyrosines, Tyr105, Tyr135, Tyr156, in Mmhk1s-z2 structurally correspond to Trp109, Tyr139, Trp161 in McpXPR. Tyr259 is located in the ‘tongue-like’ loop region. An additional aromatic residue, Tyr172, involved in hydrogen-bonding interactions with the ligand, and Asp199 are also conserved in McpX (Tyr177 and Asp208, respectively). In the amino acid sensor Mlp37 Trp154, Tyr170, and Asp277 (Trp161, Tyr177, Asp208 in McpX) are conserved but the serine ligand in Mlp37 is coordinated by charged and polar residues. In Tlp3 Tyr118, Trp150, Tyr167, Asp196 correspond to Trp109, Trp161, Tyr177, and Asp208 in McpX, but again binding of the isoleucine ligands primarily involves hydrogen-bonding and ionic interactions. The same is true for the primary amine receptor McpU in *P. putida*. In the putrescine complex of the McpUPR, the ligand is coordinated by two aspartate residues, where Asp233 is structurally equivalent to Asp208 in McpX. Once again Tyr139 (Tyr152 in McpU), Trp161 (Trp186 in McpU), Tyr177 is also conserved in McpU (Tyr202). Lastly, in the recently solved structure of the arginine-complex of the *P. aeruginosa* receptor protein PctB (PDB code: 5LT9, unpublished), Tyr101 (Trp109 in McpX), Trp128 (Trp161 in McpX), Tyr144 (Tyr177 in McpX), and Asp173 (Asp208 in McpX) are also conserved. Because only Mmhk1s-z2 shares the box-like arrangement of aromatic residues around the ligand, it is tempting to hypothesize that the unknown ligands of the former receptor might also be QACs.

The highly conserved Asp208 residue is critical for QAC binding of McpXPR

During the analysis of the ligand-binding pocket, we were struck by the conspicuous absence of interactions with charged amino acid residues in the immediate vicinity. In all other complexes, the amino acid ligands are stabilized by at least two ionic side chains from the receptor, yet in the McpXPR-proline betaine complex, the carboxylate group of the ligand is hydrogen-bonded to peptide bond amides groups, while the positive charge of the amine group is coordinated by the  $\pi$ -systems of the surrounding aromatic amino acid residues. The comparison to other receptor complexes drew our attention to the highly conserved Asp208 residue. In other structures, the equivalent residue provided pivotal hydrogen-bonding contacts with the ligand, the most recent examples being the complex of the McpU ligand-binding domain with putrescine, where the

equivalent aspartate (Asp233) interacts with a primary amine group of the ligand. However, in the McpXPR-proline betaine complex, the positively charged amine group lacks polar hydrogen atoms and the side chain of Asp208 is positioned more than 4 Å away from that nitrogen atom (Figure 5.2). To determine whether or not the charge of Asp208 is nevertheless important for balancing the positive charge of the quaternary amine group, we replaced the residue with an asparagine and purified the McpXD208N-PR variant protein. Using ITC, we determined binding isotherms of the McpXD208N-PR variant for both choline and proline betaine. Remarkably, the conservative D208N mutation completely abolished binding of McpXPR to either ligand, demonstrating that Asp208 is essential for McpX function despite being positioned more than 4 Å away from the charged amino group of the ligand (Figure 5.3). The demonstrated significance of Asp208 may also explain the broad conservation of two additional residues that, at least in McpX, do not directly interact with the ligand: Ser190 forms hydrogen bonds with Tyr177 and Asp208 thus providing important second-tier contacts for shaping the binding pocket.

Molecular docking provides rationalization of the differential binding affinity of McpXPR for its various ligands

In a prior study, the ligand-binding properties of McpX have been profiled both qualitatively and quantitatively [21]. While the protein showed significant affinity toward only one of all proteinogenic amino acids tested, namely proline ( $K_D = 45.2 \mu\text{M}$ ), it displayed broad specificity toward QACs, with choline being the tightest binding ligand ( $K_D = 138 \text{ nM}$ ) and betonicine binding with the weakest affinity ( $K_D = 2.3 \text{ mM}$ ) [21]. We can rationalize the 10-fold lower affinity of proline compared with the closely related proline betaine ( $K_D = 3.8 \mu\text{M}$ ) with the fact that the ligand-binding pocket is primarily formed by aromatic residues that create a largely hydrophobic environment in that region. Therefore, the non-polar methyl groups surrounding the charged nitrogen atom provide better complementarity within the binding pocket. The difference in affinity is even greater between proline betaine and betonicine even though the only difference is an additional hydroxyl group in the C4 position of the latter. If betonicine is overlaid onto proline betaine in our experimentally determined structure, the hydroxyl group would likely force a conformational change in the ring of the ligand to avoid clashing of the hydroxyl group with the aromatic ring of Trp161 (Figure 5.4B). Using a molecular docking protocol validated through successful re-docking of proline betaine, one can predict a reasonably well-fitting complex with

betonicine, but in that complex, the quaternary amine group of betonicine is slightly displaced. Either this displacement of the quaternary amine or a forced conformational change in the protein needed to accommodate the additional hydroxyl group could be the underlying causes for the 1000-fold reduced affinity. Additionally, we performed molecular docking studies to determine the structural basis of this selectivity not only for proline betaine (control) and betonicine but four other ligands known to bind McpXPR, including proline, trigonelline, choline, and glycine betaine. Remarkably, despite their structural variability, the quaternary amine groups of all six ligands docked within 1 Å of the position of proline betaine in the experimental model (Figure 5.4C,D and 5.7). As the quaternary amine group constitutes the single unifying feature of all six ligands, it appears that the cation- $\pi$  interactions and ionic interactions of the quaternary amine group with Asp208 are the pivotal specificity determinants. Moreover, ligand features causing the positively charged nitrogen atom to shift from its ideal position right at the center between the four coordinating  $\pi$ -systems lead to a reduced binding affinity. Therefore, choline is likely the ligand with the highest affinity because its flexible backbone structure permits ideal positioning of the quaternary amine group. As discussed below, substrate-binding proteins (SBPs) involved in betaine transport also utilize cation- $\pi$  interactions to bind their cognate ligands and the positions of the quaternary amine groups were observed to be highly conserved in the various QAC complexes [38–40]. The results of our modeling and docking studies show nice shape complementarity between McpXPR and the high-affinity ligands proline betaine, proline, choline, and glycine betaine, suggesting that they all bind in the pocket without undergoing major conformational changes or forcing changes in the protein. In the weaker binding ligands, a good fit into the ligand-binding pocket of McpXPR without major clashes is prevented by either a rigid ring structure (trigonelline,  $K_D = 88.5 \mu\text{M}$ ) or by a bulky substituent (betonicine,  $K_D = 2.3 \text{ mM}$ ).

## DISCUSSION

dCACHE\_1 domains are broadly conserved, periplasmic ligand sensing modules found in bacterial chemoreceptors and histidine kinases [23]. Generally, it is the membrane-distal CACHE domain that binds the signaling ligand, while the second CACHE domain is thought to mediate signal transmission to the membrane [1]. However, it is worth noting that, in a departure from the existing paradigm, a recently discovered lactate sensor utilizes its membrane-proximal CACHE domain to directly sense lactate [36]. Ligand-binding appears to alter the bending of the central



helix causing the membrane-proximal section of the helix to be more splayed but there also appears to be a slight rotation associated with ligand binding that further separates the C-terminal ends of the two central helices in the dimer [41–45].

Although the present structure constitutes the first example of a chemotactic sensor with specificity for small QACs, we found many protein complexes in the Protein Data Bank with bound choline, glycine betaine or proline betaine. There are four entries for proline betaine complexes in the Protein Data Bank. Three of these structures belong to SBPs that mediate the uptake of small molecules through bacterial ABC transporters. Like McpX, SBPs ProX, OpuA, OpuAC specifically target glycine betaine and proline betaine but do not facilitate amino acid transport [46]. In fact, ProX from *E. coli* is even more selective than McpX as it binds proline betaine with a KD of  $\sim 5 \mu\text{M}$ , while no binding of proline has been observed [40,47]. Remarkably, even though the SBP structures bear little resemblance to the structure of McpXPR, the binding pockets are strikingly similar. In the structure of ligand-binding protein ProX from the hyperthermophilic archaeon *Archaeoglobus fulgidus* (PDB code 1SW1 [39]), the quaternary amine of proline betaine ligand is engaged in cation– $\pi$  interactions with four tyrosines, while the carboxylate group forms ionic interactions with Arg and Lys residues [39]. This protein was also crystallized in complex with glycine betaine (PDB code 1SW2). Glycine betaine forms the same key contacts as proline betaine and the quaternary amine groups are in the same position. In the structure of the *E. coli* ProX protein, the bound proline betaine ligand is coordinated by three Trp residues, while the carboxylate group is stabilized through hydrogen-bonding contacts with two backbone peptide groups and an ionic contact with a histidine side chain [48]. Again, a glycine betaine complex formed the same contacts between protein and ligand (PDB codes 1R9L and 1R9Q) [40]. The SBP OpuAC (PDB code: 2B4M) from *Bacillus subtilis* displays the same fold as ProX [38]. Here, the proline betaine ligand is also forming cation– $\pi$  contacts with three Trp residues but the carboxylate group is stabilized only through hydrogen-bonding contacts with backbone peptide groups akin to those observed in McpX. Interestingly, at least in the case of OpuAC, the closest negatively charged amino acid residue, Glu181, is  $6.1 \text{ \AA}$  away from the quaternary amine group in the proline betaine complex suggesting that in this instance the ion–ion interactions are perhaps not as important as we observed in McpX (Figure 5.5). OpuAC was also crystallized in complex with glycine betaine (PDB code 2B4L) and the key interactions are almost identical but the absence of

the ring structure from the ligand permitted the same glutamate side chain to move within 4.6 Å of the quaternary amine group.

The present work is less focused on the mechanism of signal transduction but rather on the question of how a ligand-binding module has adapted to preferentially bind small quaternary amine-containing osmolytes such as choline and proline betaine over amino acids. How does a protein maintain a strong affinity for QACs while at the same time discriminating against binding of amino acids? The answer appears to be the creation of a ligand-binding pocket rich in aromatic amino acid residues with few polar or charged amino acids in the immediate vicinity of the ligand to accommodate the unusual combination of a positive charge surrounded by hydrophobic alkyl groups as presented in QACs. Osmoprotectants such as glycine betaine and proline betaine are excluded from the immediate surface of proteins and are known to stabilize proteins by increasing the water concentration in the immediate vicinity of the macromolecules [49]. The sparsity of aromatic amino acids on protein surfaces stands in striking contrast with their prevalence in the binding pockets of the SBPs and McpX and thus the SBP structures provided an explanation for how a receptor can evolve affinity for a ligand that generally does not interact with proteins. Collectively, the comparison of the SBP complexes with the McpXPR-proline betaine complex demonstrates a striking example of convergent evolution highlighting that cation- $\pi$  interactions are ideal for providing binding specificity toward QACs. Both tyrosine and tryptophan residues are suitable for forming the interactions. The position of the quaternary amine group is fixed within the pocket, additional stabilizing interactions are quite variable but can serve to discriminate between QAC ligands [21,48].

#### ABBREVIATIONS

ADT, AutoDock Tools; dCACHE\_1, dual CACHE domain; ITC, isothermal titration calorimetry; MCPs, methyl-accepting chemotaxis proteins; McpXPR, McpX ligand-binding periplasmic region; PDB, Protein Data Bank; PMSF, phenylmethylsulfonyl fluoride; QACs, quaternary ammonium compounds; RMSD, root-mean-square deviation; SBPs, substrate-binding proteins; TCEP, tri(2-carboxyethyl)phosphine.

#### AUTHOR CONTRIBUTION

M.S. performed the structural studies and analysis, K.K.C. isolated proteins and performed the ITC studies, J.M. M. obtained the initial crystals for McpX, assisted with the model building and manuscript, B.A.W. developed protocols for McpX purification and characterization, A.M.B. performed the molecular docking studies, B.E.S. co-wrote the manuscript, provided funding for the project and supervised the biochemical studies, F.D.S. supervised all structural and biochemical analysis, co-wrote the manuscript, and provided funding for the project.

#### FUNDING

The present study was supported by National Science Foundation grant MCB-1253234 to B.E.S and MCB-1817652 to B.E.S and F.D.S. Use of the Advanced Photon Source was supported by the U. S. Department of Energy, Office of Science, Office of Basic Energy Sciences, under Contract No. W-31-109-Eng-38.

#### COMPETING INTERESTS

The Authors declare that there are no competing interests associated with the manuscript.

## REFERENCES

1. Ortega, A., Zhulin, I.B. and Krell, T. (2017) Sensory repertoire of bacterial chemoreceptors. *Microbiol. Mol. Biol. Rev.* 81, 1–28 <https://doi.org/10.1128/MMBR.00033-17>
2. Matilla, M.A. and Krell, T. (2018) The effect of bacterial chemotaxis on host infection and pathogenicity. *FEMS Microbiol. Rev.* 42, 40–67 <https://doi.org/10.1093/femsre/fux052>
3. Scharf, B.E., Hynes, M.F. and Alexandre, G.M. (2016) Chemotaxis signaling systems in model beneficial plant-bacteria associations. *Plant Mol. Biol.* 90, 549–559 <https://doi.org/10.1007/s11103-016-0432-4>
4. Hazelbauer, G.L. (2012) Bacterial chemotaxis: the early years of molecular studies. *Annu. Rev. Microbiol.* 66, 285–303 <https://doi.org/10.1146/annurev-micro-092611-150120>
5. Parkinson, J.S., Hazelbauer, G.L. and Falke, J.J. (2015) Signaling and sensory adaptation in *Escherichia coli* chemoreceptors: 2015 update. *Trends Microbiol.* 23, 257–266 <https://doi.org/10.1016/j.tim.2015.03.003>
6. Hazelbauer, G.L., Falke, J.J. and Parkinson, J.S. (2008) Bacterial chemoreceptors: high-performance signaling in networked arrays. *Trends Biochem. Sci.* 33, 9–19 <https://doi.org/10.1016/j.tibs.2007.09.014>
7. Berg, H.C. (2003) The rotary motor of bacterial flagella. *Annu. Rev. Biochem.* 72, 19–54 <https://doi.org/10.1146/annurev.biochem.72.121801.161737>
8. Scharf, B. (2002) Real-time imaging of fluorescent flagellar filaments of *Rhizobium lupini* H13-3: flagellar rotation and pH-induced polymorphic transitions. *J. Bacteriol.* 184, 5979–5986 <https://doi.org/10.1128/JB.184.21.5979-5986.2002>
9. Maddock, J.R. and Shapiro, L. (1993) Polar location of the chemoreceptor complex in the *Escherichia coli* cell. *Science* 259, 1717–1723 <https://doi.org/10.1126/science.8456299>
10. Sourjik, V. and Berg, H.C. (2002) Receptor sensitivity in bacterial chemotaxis. *Proc. Natl Acad. Sci. U.S.A.* 99, 123–127 <https://doi.org/10.1073/pnas.011589998>
11. Sourjik, V. and Berg, H.C. (2004) Functional interactions between receptors in bacterial chemotaxis. *Nature* 428, 437–441 <https://doi.org/10.1038/nature02406>
12. Colin, R. and Sourjik, V. (2017) Emergent properties of bacterial chemotaxis pathway. *Curr. Opin. Microbiol.* 39, 24–33 <https://doi.org/10.1016/j.mib.2017.07.004>
13. Gulash, M., Ames, P., Larosiliere, R.C. and Bergman, K. (1984) *Rhizobia* are attracted to localized sites on legume roots. *Appl. Environ. Microbiol.* 48, 149–152
14. Soby, S. and Bergman, K. (1983) Motility and chemotaxis of *Rhizobium meliloti* in soil. *Appl. Environ. Microbiol.* 46, 995–998
15. Caetano-Anolles, G., Wall, L.G., De Micheli, A.T., Macchi, E.M., Bauer, W.D. and Favelukes, G. (1988) Role of motility and chemotaxis in efficiency of nodulation by *Rhizobium meliloti*. *Plant Physiol.* 86, 1228–1235 <https://doi.org/10.1104/pp.86.4.1228>

16. Meier, V.M., Muschler, P. and Scharf, B.E. (2007) Functional analysis of nine putative chemoreceptor proteins in *Sinorhizobium meliloti*. *J. Bacteriol.* 189, 1816–1826 <https://doi.org/10.1128/JB.00883-06>
17. Meier, V.M. and Scharf, B.E. (2009) Cellular localization of predicted transmembrane and soluble chemoreceptors in *Sinorhizobium meliloti*. *J. Bacteriol.* 191, 5724–5733 <https://doi.org/10.1128/JB.01286-08>
18. Webb, B.A., Compton, K.K., Del Campo, J.S.M., Taylor, D., Sobrado, P. and Scharf, B.E. (2017) *Sinorhizobium meliloti* chemotaxis to multiple amino acids is mediated by the chemoreceptor McpU. *Mol. Plant Microbe Interact.* 30, 770–777 <https://doi.org/10.1094/MPMI-04-17-0096-R>
19. Webb, B.A., Helm, R.F. and Scharf, B.E. (2016) Contribution of individual chemoreceptors to *Sinorhizobium meliloti* chemotaxis towards amino acids of host and nonhost seed exudates. *Mol. Plant Microbe Interact.* 29, 231–239 <https://doi.org/10.1094/MPMI-12-15-0264-R>
20. Webb, B.A., Hildreth, S., Helm, R.F. and Scharf, B.E. (2014) *Sinorhizobium meliloti* chemoreceptor McpU mediates chemotaxis toward host plant exudates through direct proline sensing. *Appl. Environ. Microbiol.* 80, 3404–3415 <https://doi.org/10.1128/AEM.00115-14>
21. Webb, B.A., Compton, K.K., Castañeda Saldaña, R., Arapov, T., Ray, W.K., Helm, R.F. et al. (2017) *Sinorhizobium meliloti* chemotaxis to quaternary ammonium compounds is mediated by the chemoreceptor McpX. *Mol. Microbiol.* 103, 333–346 <https://doi.org/10.1111/mmi.13561>
22. Welsh, D.T. (2000) Ecological significance of compatible solute accumulation by microorganisms: from single cells to global climate. *FEMS Microbiol. Rev.* 24, 263–290 <https://doi.org/10.1111/j.1574-6976.2000.tb00542.x>
23. Upadhyay, A.A., Fleetwood, A.D., Adebali, O., Finn, R.D. and Zhulin, I.B. (2016) Cache domains that are homologous to, but different from PAS domains comprise the largest superfamily of extracellular sensors in prokaryotes. *PLoS Comput. Biol.* 12, e1004862 <https://doi.org/10.1371/journal.pcbi.1004862>
24. Zhulin, I.B., Nikolskaya, A.N. and Galperin, M.Y. (2003) Common extracellular sensory domains in transmembrane receptors for diverse signal transduction pathways in bacteria and archaea. *J. Bacteriol.* 185, 285–294 <https://doi.org/10.1128/JB.185.1.285-294.2003>
25. Anantharaman, V. and Aravind, L. (2000) Cache - a signaling domain common to animal Ca(2+)-channel subunits and a class of prokaryotic chemotaxis receptors. *Trends Biochem. Sci.* 25, 535–537 [https://doi.org/10.1016/S0968-0004\(00\)01672-8](https://doi.org/10.1016/S0968-0004(00)01672-8)
26. Anantharaman, V., Koonin, E.V. and Aravind, L. (2001) Regulatory potential, phyletic distribution and evolution of ancient, intracellular small-molecule-binding domains. *J. Mol. Biol.* 307, 1271–1292 <https://doi.org/10.1006/jmbi.2001.4508>
27. Adams, P.D., Afonine, P.V., Bunkóczi, G., Chen, V.B., Davis, I.W., Echols, N. et al. (2010) PHENIX: a comprehensive Python-based system for macromolecular structure solution. *Acta Crystallogr. D Biol. Crystallogr.* 66(Pt 2), 213–221 <https://doi.org/10.1107/S0907444909052925>

28. Karplus, P.A. and Diederichs, K. (2015) Assessing and maximizing data quality in macromolecular crystallography. *Curr. Opin. Struct. Biol.* 34, 60–68 <https://doi.org/10.1016/j.sbi.2015.07.003>
29. Morris, G.M., Huey, R., Lindstrom, W., Sanner, M.F., Belew, R.K., Goodsell, D.S. et al. (2009) Autodock4 and AutoDockTools4: automated docking with selective receptor flexibility. *J. Comput. Chem.* 30, 2785–2791 <https://doi.org/10.1002/jcc.21256>
30. Trott, O. and Olson, A.J. (2010) Autodock Vina: improving the speed and accuracy of docking with a new scoring function, efficient optimization, and multithreading. *J. Comput. Chem.* 31, 455–461 PMID:19499576
31. Zhang, Z. and Hendrickson, W.A. (2010) Structural characterization of the predominant family of histidine kinase sensor domains. *J. Mol. Biol.* 400, 335–353 <https://doi.org/10.1016/j.jmb.2010.04.049>
32. Liu, Y.C., Machuca, M.A., Beckham, S.A., Gunzburg, M.J. and Roujeinikova, A. (2015) Structural basis for amino-acid recognition and transmembrane signalling by tandem Per-Arnt-Sim (tandem PAS) chemoreceptor sensory domains. *Acta Crystallogr. D Biol. Crystallogr.* 71(Pt 10), 2127–2136 <https://doi.org/10.1107/S139900471501384X>
33. Nishiyama, S.-., Takahashi, Y., Yamamoto, K., Suzuki, D., Itoh, Y., Sumita, K. et al. (2016) Identification of a *Vibrio cholerae* chemoreceptor that senses taurine and amino acids as attractants. *Sci. Rep.* 6, 20866 <https://doi.org/10.1038/srep20866>
34. Glekas, G.D., Mulhern, B.J., Kroc, A., Duelfer, K.A., Lei, V., Rao, C.V. et al. (2012) The *Bacillus subtilis* chemoreceptor McpC senses multiple ligands using two discrete mechanisms. *J. Biol. Chem.* 287, 39412–39418 <https://doi.org/10.1074/jbc.M112.413518>
35. Rico-Jiménez, M., Muñoz-Martínez, F., Garcia-Fontana, C., Fernandez, M., Morel, B., Ortega, A. et al. (2013) Paralogous chemoreceptors mediate chemotaxis towards protein amino acids and the non-protein amino acid gamma-aminobutyrate (GABA). *Mol. Microbiol.* 88, 1230–1243 <https://doi.org/10.1111/mmi.12255>
36. Machuca, M.A., Johnson, K.S., Liu, Y.C., Steer, D.L., Ottemann, K.M. and Roujeinikova, A. (2017) *Helicobacter pylori* chemoreceptor TlpC mediates chemotaxis to lactate. *Sci. Rep.* 7, 14089 <https://doi.org/10.1038/s41598-017-14372-2>
37. Gavira, J.A., Ortega, A., Martín-Mora, D., Conejero-Muriel, M.T., Corral-Lugo, A., Morel, B. et al. (2018) Structural basis for polyamine binding at the dCACHE domain of the McpU chemoreceptor from *Pseudomonas putida*. *J. Mol. Biol.* 430, 1950–1963 <https://doi.org/10.1016/j.jmb.2018.05.008>
38. Horn, C., Sohn-Bösser, L., Breed, J., Welte, W., Schmitt, L. and Bremer, E. (2006) Molecular determinants for substrate specificity of the ligand-binding protein OpuAC from *Bacillus subtilis* for the compatible solutes glycine betaine and proline betaine. *J. Mol. Biol.* 357, 592–606 <https://doi.org/10.1016/j.jmb.2005.12.085>
39. Schiefner, A., Holtmann, G., Diederichs, K., Welte, W. and Bremer, E. (2004) Structural basis for the binding of compatible solutes by ProX from the hyperthermophilic archaeon

Archaeoglobus fulgidus. J. Biol. Chem. 279, 48270–48281  
<https://doi.org/10.1074/jbc.M403540200>

40. Schiefner, A., Breed, J., Bösser, L., Kneip, S., Gade, J., Holtmann, G. et al. (2004) Cation- $\pi$  interactions as determinants for binding of the compatible solutes glycine betaine and proline betaine by the periplasmic ligand-binding protein ProX from *Escherichia coli*. J. Biol. Chem. 279, 5588–5596 <https://doi.org/10.1074/jbc.M309771200>

41. Ames, P., Hunter, S. and Parkinson, J.S. (2016) Evidence for a helix-clutch mechanism of transmembrane signaling in a bacterial chemoreceptor. J. Mol. Biol. 428, 3776–3788 <https://doi.org/10.1016/j.jmb.2016.03.017>

42. Ottemann, K.M., Xiao, W., Shin, Y.K. and Koshland, Jr., D.E. (1999) A piston model for transmembrane signaling of the aspartate receptor. Science 285, 1751–1754 <https://doi.org/10.1126/science.285.5434.1751>

43. Yu, D., Ma, X., Tu, Y. and Lai, L. (2015) Both piston-like and rotational motions are present in bacterial chemoreceptor signaling. Sci. Rep. 5, 8640 <https://doi.org/10.1038/srep08640>

44. Chervitz, S.A. and Falke, J.J. (1996) Molecular mechanism of transmembrane signaling by the aspartate receptor: a model. Proc. Natl Acad. Sci. U.S.A. 93, 2545–2550 <https://doi.org/10.1073/pnas.93.6.2545>

45. Gushchin, I., Melnikov, I., Polovinkin, V., Ishchenko, A., Yuzhakova, A., Buslaev, P. et al. (2017) Mechanism of transmembrane signaling by sensor histidine kinases. Science 356, eaah6345 <https://doi.org/10.1126/science.aah6345>

46. Haardt, M., Kempf, B., Faatz, E. and Bremer, E. (1995) The osmoprotectant proline betaine is a major substrate for the binding-protein-dependent transport system ProU of *Escherichia coli* K-12. Mol. Gen. Genet. 246, 783–786 <https://doi.org/10.1007/BF00290728>

47. Barron, A., Jung, J.U. and Villarejo, M. (1987) Purification and characterization of a glycine betaine binding protein from *Escherichia coli*. J. Biol. Chem. 262, 11841–6

48. Smits, S.H., Hoing, M., Lecher, J., Jebbar, M., Schmitt, L. and Bremer, E. (2008) The compatible-solute-binding protein OpuAC from *Bacillus subtilis*: ligand binding, site-directed mutagenesis, and crystallographic studies. J. Bacteriol. 190, 5663–5671 <https://doi.org/10.1128/JB.00346-08>

49. Horn, C., Jenewein, S., Sohn-Bösser, L., Bremer, E. and Schmitt, L. (2005) Biochemical and structural analysis of the *Bacillus subtilis* ABC transporter OpuA and its isolated subunits. J. Mol. Microbiol. Biotechnol. 10, 76–91 <https://doi.org/10.1159/000091556>

50. Karplus, P.A. and Diederichs, K. (2012) Linking crystallographic model and data quality. Science 336, 1030–1033 <https://doi.org/10.1126/science.1218231>

51. Brünger, A.T. (1992) Free R value: a novel statistical quantity for assessing the accuracy of crystal structures. Nature 355, 472–475 <https://doi.org/10.1038/355472a0>

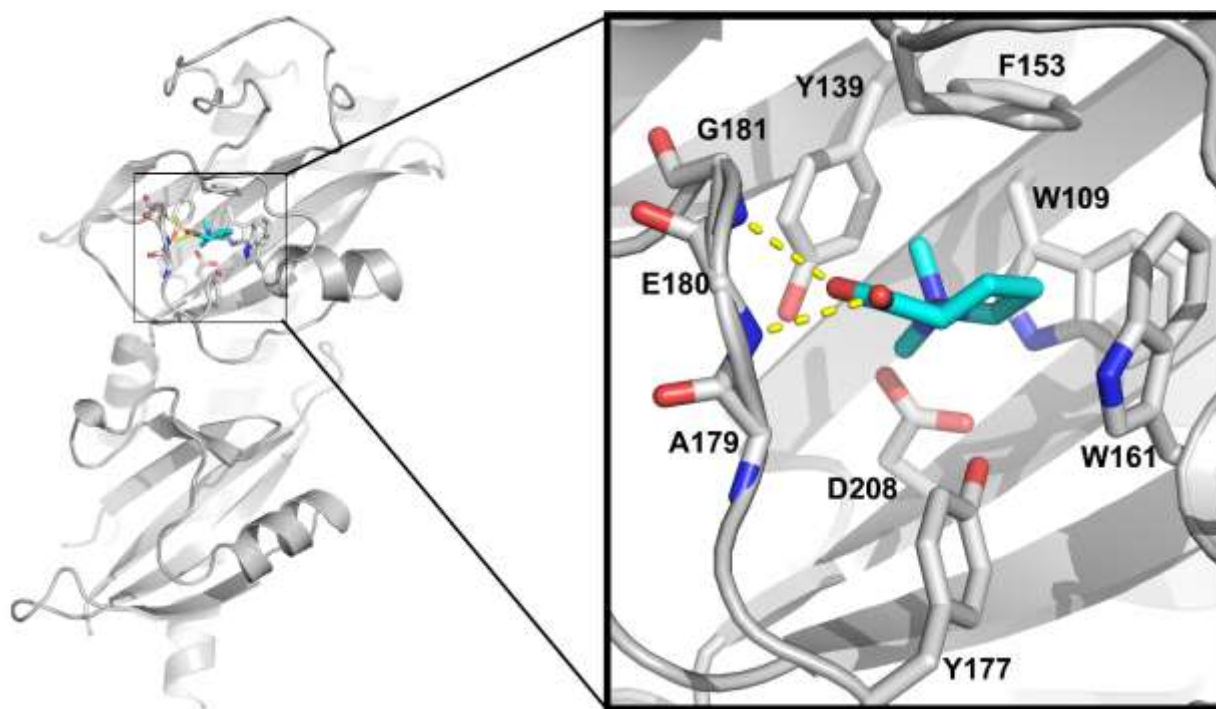
**Table 1. Data collection and refinement statistics.**

	<b>McpX<sup>PR</sup></b>
Wavelength (Å)	0.979
Resolution range (Å)	78.74 - 2.7 (2.83 - 2.7)
Space group	P 6 <sub>2</sub> 2 2
Unit cell (Å)	119.04 119.04 121.58
Total reflections	101,395 (13,428)
Unique reflections	14,450 (1877)
Multiplicity	7.0 (7.2)
Completeness (%)	99.7 (100)
Mean I/sigma(I)	9.4 (1.7)
Wilson B-factor	77.10
R-merge	0.072 (0.766)
R-meas	0.080 (0.835)
CC1/2	0.998 (0.83)
Reflections used in refinement	14,121 (1,348)
Reflections used for R-free	704 (64)
R-work	0.220 (0.3895)
R-free	0.2510 (0.3934)
CC(work)	0.945 (0.494)
CC(free)	0.951 (0.419)
Number of non-hydrogen atoms	2032
Macromolecules	2012
Ligands	18
Protein residues	269
RMS bonds (Å)	0.004
RMS angles (°)	0.60
Ramachandran favored (%)	93.63
Ramachandran allowed (%)	6.37
Average B-factor (Å <sup>2</sup> )	100.6
Macromolecules	100.42
Ligands	121.53

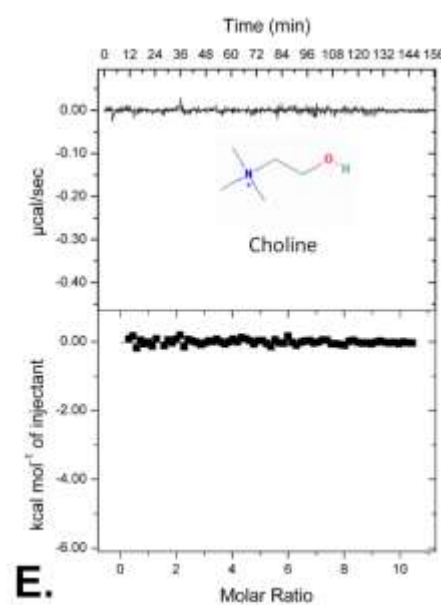
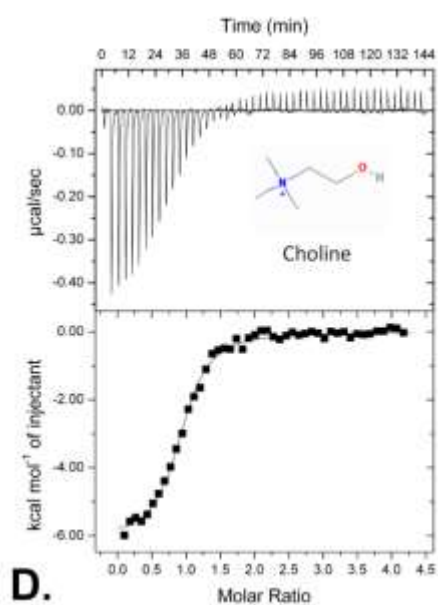
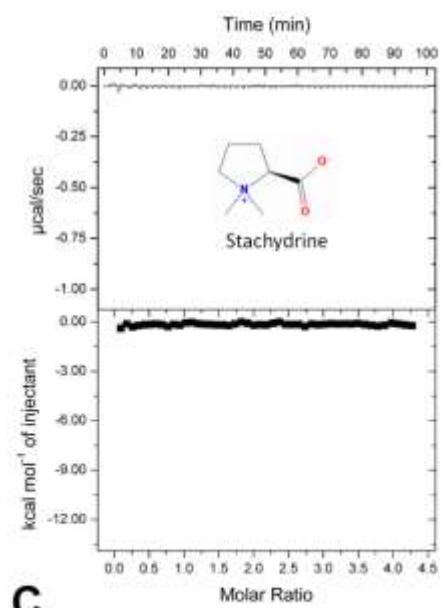
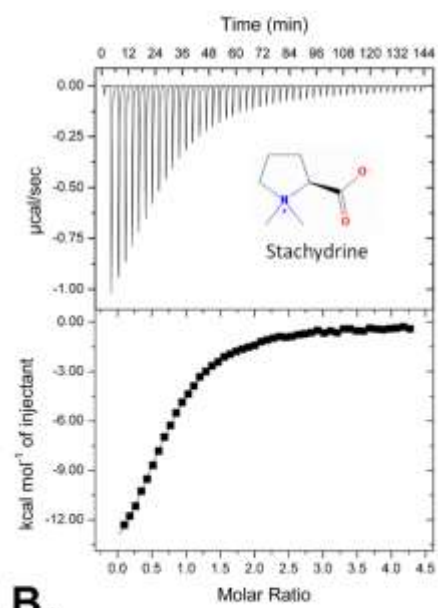
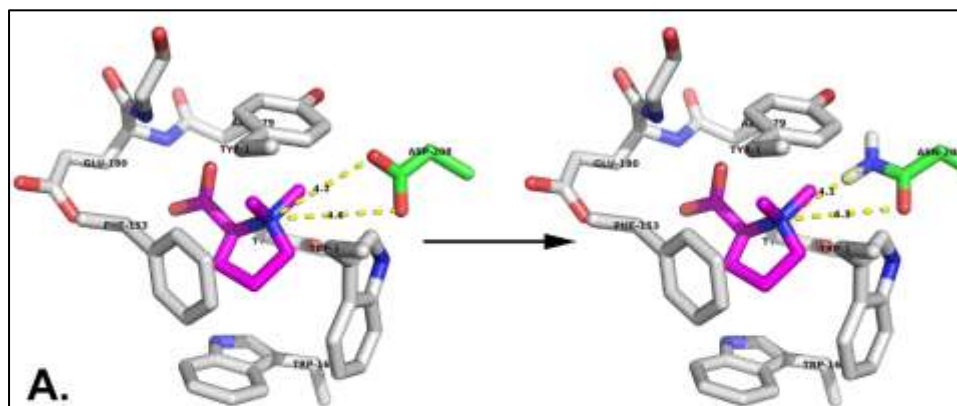
The values in parentheses relate to the highest resolution shell from 2.72 – 2.7 Å.  $R_{\text{merge}} = \frac{\sum ||I - \langle I \rangle|}{\sum I}$ , where  $I$  is the observed intensity and  $\langle I \rangle$  is the average intensity obtained from multiple observations of symmetry-related reflections after the rejection of significant outliers.  $CC_{1/2}$  = Pearson correlation coefficient between random half-datasets.  $R\text{-work} = \frac{\sum ||F_o| - |F_c||}{\sum |F_o|}$ , R-free defined by Brunger (1992)



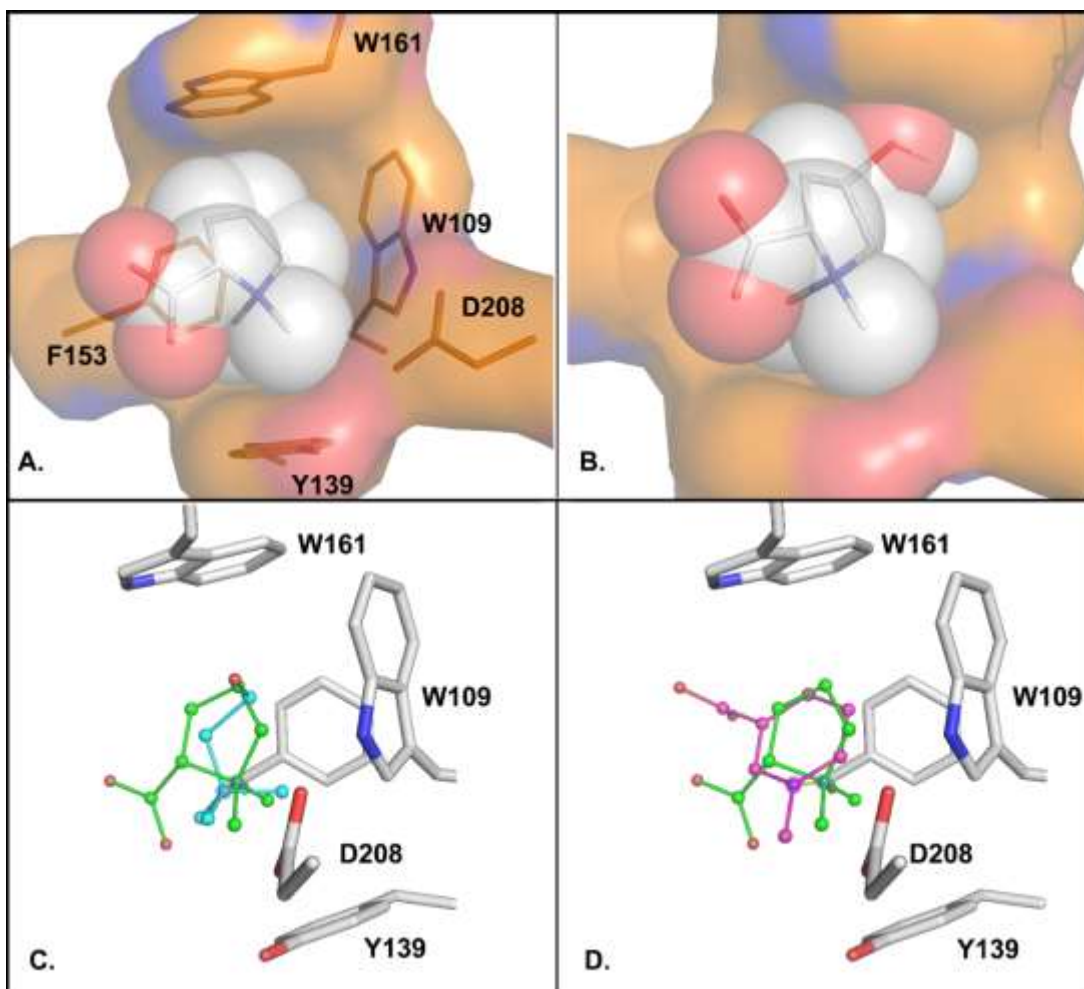




**Fig. 5.2** Close-up view of the proline betaine binding pocket of McpX<sup>PR</sup>. Particularly noteworthy is prevalence of interacting aromatic residues that box-in the ligand on four sides. The dashed lines mark the  $\sim 2.9$  Å distance between the hydrogen-bonded amide groups from the protein and the carboxylate group of the ligand.

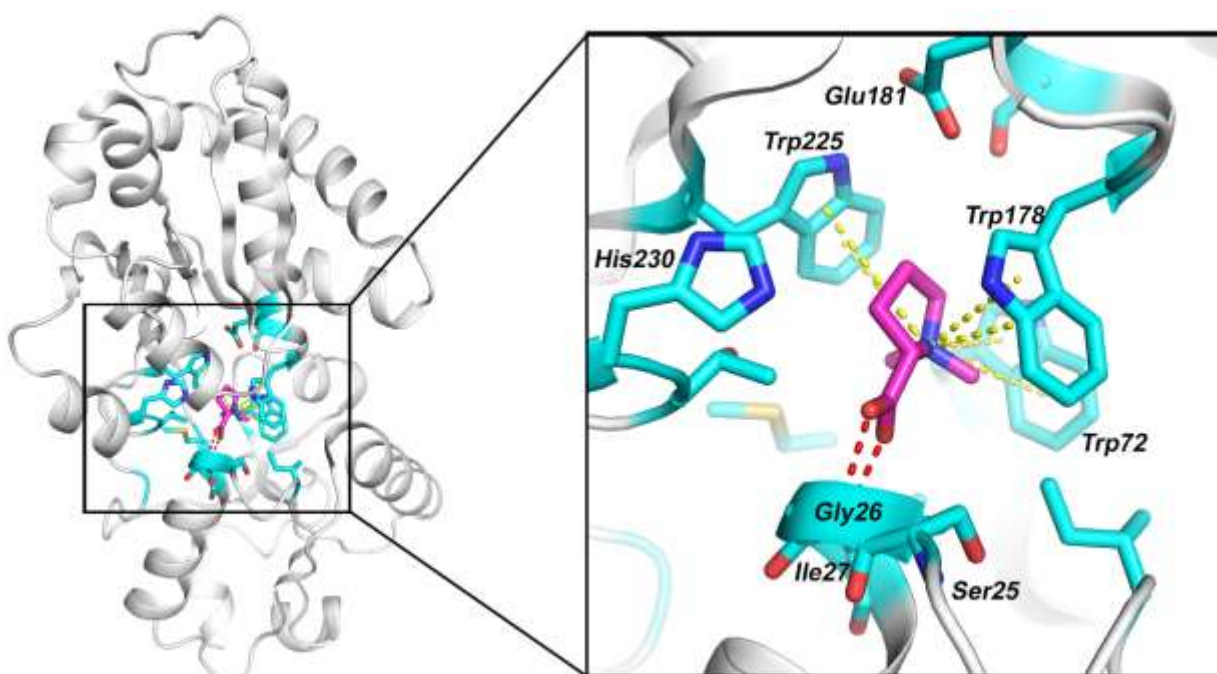


**Fig. 5.3.** (A) Structural consequences of the D208N mutation in McpX<sup>PR</sup>. (B-E) Results of ITC binding studies with either the McpX<sup>PR</sup> protein or the McpX<sup>D208N-PR</sup> variant. Upper and lower panels show the raw titration data and the isotherms derived by integrating peaks from the raw data, respectively. (B) Titration of McpX<sup>PR</sup> with proline betaine. (C) Titration of McpX<sup>D208N-PR</sup> with proline betaine. (D) Titration of McpX<sup>PR</sup> with choline. (E) Titration of McpX<sup>D208N-PR</sup> with choline.

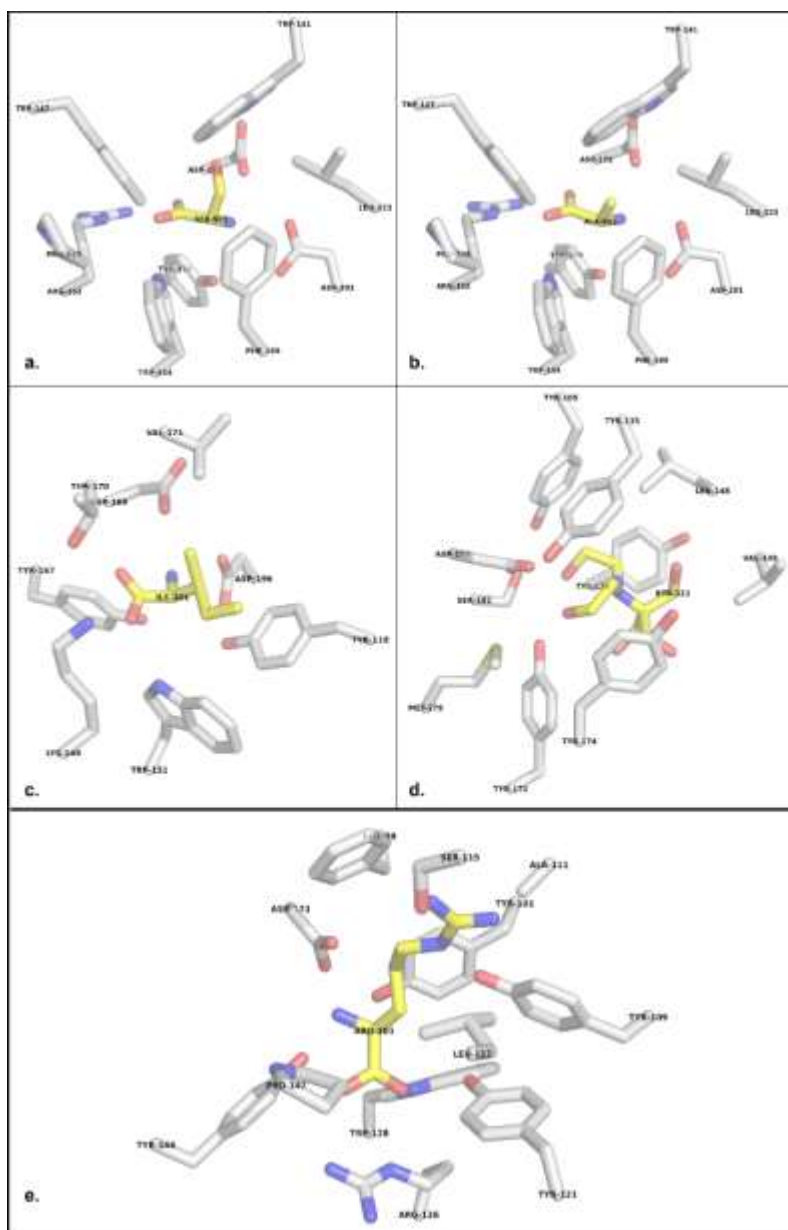


**Fig. 5.4** Possible poses of the various McpX<sup>PR</sup> ligands. (A) Spacefill showing the fit of the proline betaine ligand in the experimentally determined structure. Proline is expected to bind similarly, with the two methyl groups on the nitrogen being replaced with hydrogen atoms. (B) Space fill showing the fit of betonicine in McpX<sup>PR</sup> created by adding a hydroxyl group to the experimentally fit proline betaine to demonstrate the clash of the hydroxyl group with binding site residues to explain why betonicine is bound with significantly lower affinity. (C) Lowest energy pose for a docked of choline McpX<sup>PR</sup> complex (cyan). Proline betaine (green) is included for comparison. The two quaternary amino groups are separated by 0.6 Å. (D) Lowest energy pose obtained from docking trigonelline (hot pink) into the McpX<sup>PR</sup> structure. Proline betaine (green) is included for comparison. The two quaternary amino groups are separated by 0.8 Å.

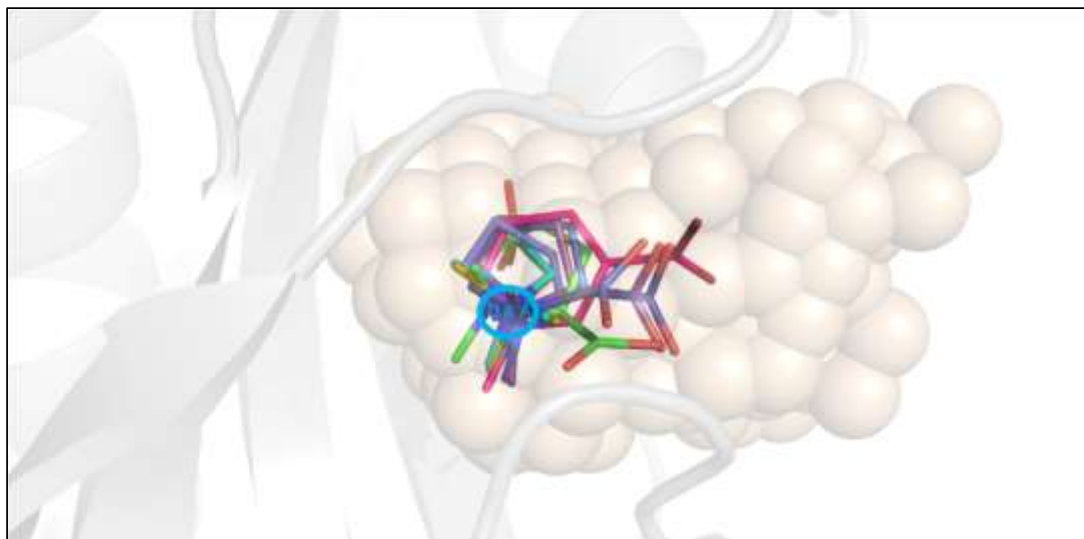




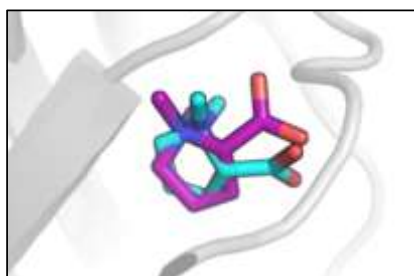
**Fig. 5.5** Overall structure and close-up view of the proline betaine binding pocket of the SBP OpuAC. Noteworthy, are the similarities between the protein-ligand interactions in this structure and those observed in the McpX<sup>PR</sup>-proline betaine complex.



**Fig. 5.6 Binding site geometries found in the dCACHE regions of other MCPs:** Shown are the binding sites of Mlp37 from *Vibrio cholerae* (PDB code: 5avf and 3c8c respectively, (A and B)), as the isoleucine ligand in the *Campylobacter jejuni* chemoreceptor Tlp3 (PDB code: 4xmr, (C)), as Bis-Tris in the putative histidine kinase mmhk1s-z2 from *Methanosarcina frisia* (PDB code: 3lia, (D)), and as the arginine-ligand in the *P. aeruginosa* receptor protein PctB (pdb code: 5LT9, unpublished) (E).



**A.**



**B.**

**Fig. 5.7. A.** Overlay of experimental position of proline betaine (shown as lines, green) with lowest energy docked poses of proline betaine (purple), betonicine (dark blue), choline (cyan), trigonelline (hot pink), and glycine betaine (yellow). The McpX<sup>PR</sup> structure is shown as grey cartoon, with tan spheres representing the binding pocket volume, as calculated by Metapocket 2.0. The furthestmost distance between ligand nitrogen atoms is 0.8 Å. **B.** Superposition of the proline betaine modelled into the experimentally determined electron density in the McpX<sup>PR</sup> crystal structure (purple sticks) and the re-docked proline betaine pose (cyan sticks). The shown re-docked ligand represents the lowest energy pose with a predicted binding energy of -6.2 kcal/mol. The RMSD calculated between the crystallized ligand and the re-docked ligand was 1.1 Å. The quaternary amine groups are in essentially identical positions



## Chapter 6 – Final Discussion

A panoply of genetic, phenotypic, biochemical, and structural data has given us a rigorous understanding of the sensors bacteria use to navigate their environments. In analogy to the human sense, the receptors assemble into a central structure like a “nose”. Signals travel through the bacterium in a phospho-relay akin to the electrochemical signals sent between the olfactory sensors and the brain. Just as humans only sense a specific set of compounds, the sensory profile of a bacterium is largely determined by its MCPs. *Sinorhizobium meliloti* has thus far been shown to be able to detect amino acids, QACs, and carboxylates of specific sizes using its MCP repertoire (1-4). Despite a previously established consensus, the evidence in chapter 2 indicates flavonoids are not the attractants they were once thought to be (5-8). With the binding profiles Mcp T, U, V, and X identified, four receptors remain to be characterized. McpY and IcpA are theorized to be involved in sensing the redox state of the cell or local oxygen concentrations, but this has yet to be directly proven (9, 10). The ligand profiles of McpW and McpZ are unknown, but we are not without clues to their potential function. Strains lacking *mcpZ* are deficient in chemotaxis to seed exudate of the non-host legume *Medicago arabica*, but not *M. sativa*, suggesting that its ligands are to be found in the exudate of *M. arabica* seeds (11).

The gold standard for bacterial chemotaxis experiments is the capillary assay originally developed by Julius Adler (12). This experiment tests the response of a population of bacteria to a gradient of a putative attractant or repellent. When done consistently, it gives quantitative data that allows for the ranking of attractants in terms of their strength (how many cells the attractant recruits). Comparing previous studies with chapters 3 and 4 demonstrates that the amino acids and QACs are the strongest classes of attractant; drawing up to  $10^6$  cells, followed by the 2 – 4 C monocarboxylates; attracting only  $10^5$  cells, leaving the dicarboxylates last; accruing about  $5 \times 10^4$  cells (1-4). Other compounds such as saccharides, polyamines, and aromatic acids are also attractants, but their sensors have yet to be elucidated (9). The mechanism of chemotactic sensing is understood primarily through *in-vitro* direct-binding studies, such as Isothermal Titration Calorimetry (ITC). ITC is a robust method because it directly measures interactions and allows for the determination of thermodynamic properties and the dissociation constant,  $K_d$ , which is a measure of affinity or tightness of binding (13, 14). For most MCPs, the  $K_d$  is in the tens of micromolar range; low enough to detect relevant concentrations of an attractant, but not too low

as to prevent dissociation when the cell enters an area of lower ligand concentration (2, 4, 15). There are certainly exceptions, though, with affinities as strong as 138 nM (McpX-choline), and as weak as 8.7 mM (McpV-formate). The work characterizing McpV's function in chapter 3 shows a correlation between ligand affinity and concentration of peak attraction. It appears that the ligand-receptor  $K_d$  (in molarity) is consistently 100-times lower than the peak concentration in the test capillary (though because of diffusion, the concentration sensed by the population is always lower) (4, 16). This pattern was also observed in a FRET assay measuring signal output (17). Contrastingly, ligand affinities of McpX in *S. meliloti* and McpS in *Pseudomonas aeruginosa* do not correlate with strength or peak concentration of attraction (2, 18). Therefore,  $K_d$  can only be used to make behavioral predictions in specific receptor systems. ITC can also permit a quantitation of the effects of point mutations. In chapter 3 we showed that a Y143A variant of McpV<sup>PR</sup> exhibited a reduction in propionate binding by a factor of 1,000 compared to the wild-type protein. This was corroborated when a strain harboring the mutation did not respond to the peak concentration of propionate. The same approach was used in chapter 5 to assess the importance of Asp-208, which was the only charged residue in the binding pocket of McpX<sup>PR</sup>. By mutating it to Asn, thus changing the charge but not the size of the residue, binding to stachydrine and choline was entirely abolished. Without the negative charge of Asp-208, the cage of  $\pi$ -electrons formed from Trp-109, Tyr-139, Phe-153, and Trp-161 was insufficient to bind the quaternary amine group of its ligands.

Identifying attractants can be a challenge of its own. Currently, our lab employs two approaches for connecting MCPs to their respective ligands. The first is an *in-vitro* approach that screens purified protein against chemical libraries. In this way, several dozen compounds can be tested at once, and a ligand profile can be gleaned from the pattern of positive hits. This powerful method was used in chapters 3 and 4 to identify the ligands of McpV and McpT, albeit with different levels of success. The second approach involves identifying attractants from ecological sources. In the case of *S. meliloti*, this means root and seed exudates. The QACs were originally identified as being abundant in the exudates of germinating alfalfa and spotted medic (*M. arabica*) seeds. Once a class of attractants is found, the cognate sensor needs to be determined. Several approaches can be used, but the investigations of McpX utilized a chemotaxis drop assay and single *mcp* gene deletion strains to determine that *mcpX* was most likely the QAC sensor (2). Unlike the capillary assay, this method gives qualitative data, but has the advantage of being much higher throughput.

Recently, a new method for identifying chemoreceptor function has been developed using protein chimeras. A strain of *E. coli* is made to express the ligand binding domain (LBD) of a receptor fused to the sensor kinase NarQ. Ligand binding is observed through a  $\beta$ -galactosidase reporter (19). This method is an excellent alternative when recalcitrant proteins cannot be expressed or purified for *in-vitro* studies.

In addition to giving clues about chemoreceptor specificity, finding attractants in host exudates can grant understanding on how host-symbiont sensing has evolved. Carboxylates, especially citrate and malate, are commonly found in many plant root exudates as well as in the less hospitable bulk soil (20-23). In chapter 4 we tentatively identified citrate, malate, and succinate in *S. meliloti* seed exudates, implicating a role for McpT in host seed sensing. Among the McpV ligands, only about 20 ng of glycolate per seed was detected in seed exudates, although acetate and propionate were quantified in root exudate of the legume genera *Lupinus* (4, 21). Analysis of QACs in *M. sativa* and *M. arabica* found a total 249 ng and 221 ng per seed, respectively (2). The sum of proteogenic and non-proteogenic amino acids in alfalfa seeds is about 5  $\mu$ g (3). Clearly, the most intense signal the bacteria can detect from the host would be the amino acids. However, nature of the response to each of these attractants depends on the sensory and signaling system. As demonstrated in capillary chemotaxis assays, QACs and amino acids are the strongest two classes of attractants. But what explains the vast difference in response between the carboxylates and nitrogen-containing attractants? Ligand affinity cannot be the determinant of response because most MCPs have a somewhat similar  $K_d$  range (24). A study quantifying the amounts of MCPs per cell found that 70 % of the receptor pool was made up of McpV (about 300 molecules per cell), while McpU and McpX accounted for only about 10 % of the total each. The relationship between signal strength and receptor abundance is clearly something of a paradox. On the other hand, perhaps the exceptionally low abundance of McpT found in the same study might explain why its ligands are such poor attractants (25). The other possible source for the disparity in attraction is inside the cell. Since the attractant/repellent signal travels through the signaling domain of the chemoreceptors to the base plate of CheW and CheA, structural differences in the MCPs could alter the affinity and interactions with CheW and CheA, making different receptors produce different strengths of signal given the same degree of receptor occupancy.

Thus, because amino acids are the most abundant known chemotactic signal in seed exudates, and because they elicit the strongest response from the bacterium, the amino acids are the most important class of compounds known to the recruitment of *S. meliloti* to its nascent host. Indeed, the amino acids alone account for approximately 23 % of the chemotaxis response (11). This should not be an excuse to ignore the other attractants, however, as each class of compounds is exuded at a high enough concentration to elicit chemotaxis at the surface of the seed (1-4). In addition, the QACs support plant and microbe survival as carbon and nitrogen sources, but also as compatible solutes. This type of functional molecule accumulates in drought-stressed organisms to protect from the deleterious effects of water loss or salt stress (26). The flavonoids are not attractants like once thought, but still perform important roles. Cognate rhizobial populations can be primed for symbiosis by seed-borne flavonoids before the seedling is ready for nodulation, and flavonoids promote the growth of rhizobia, but antagonize other species (27, 28). Carboxylates are simple carbon sources that the bacterium readily utilizes for energy via the TCA cycle (29-31). Furthermore, because it is a catabolite repressor of sugar metabolism, the dicarboxylate succinate is a preferred carbon source for *S. meliloti* (32, 33).

This current body of knowledge we have gleaned from seed exudates is powerful but does not address the nature of root exudate sensing or chemotaxis in the rhizosphere of mature plants. The rhizosphere is the volume surrounding a plant's roots that is directly affected by root exudation. Rhizosphere soils are much richer in carbon and nitrogen because of plant-based inputs and microbial turnover (22, 34, 35). The exudation of organic compounds happens primarily at the tips of growing plant roots called the elongation zone, as well as at root hairs (36-39). In front of the elongation zone is the root meristem, where most of the cell division occurs. This cluster of cells is protected by a sheath of dead cells called a root cap, secretion-active border cells, and lubricating mucilage. This area is rich in exudates and detritus that microorganisms feed on (35, 40, 41). Thus, the emerging volume created by the caravanning meristem is a prime new niche that will be quickly occupied by soil microbes (42). Perhaps invasion of the viscous space around meristems is what necessitates the complex, rigid flagella of rhizobacteria (43). Most compounds exuded from germinating seeds are leached from the seed coat – reservoirs which are originally deposited by the mother plant and are therefore fixed until imbibition and germination (44). The situation of root exudates is much more complex because the chemical profile varies with time (age, maturity stage, season) and status (diseased, fruiting, senescing, stressed) (39). Conceivably, any given

compound could be found at some concentration from some plant in some situation. The following are just a few examples available in literature. Carboxylate exudation changes with the flowering and fruiting periods in leguminous forbs (21). Alfalfa releases formononetin under copper (oxidative) stress (45). Citrate is widely used by plants to subvert aluminum stress by chelation (46-48). Root amino acid efflux is influenced by fungal and bacterial neighbors through secondary metabolites (49). Maize seedlings release more ribitol and glucose when deprived of iron (50). One caveat of the above sources is that most are performed under sterile culture conditions. Of course, this is done to quantify plant excretion without interference from microbes that would normally consume these compounds. However, no plant exists naturally in a sterile environment, so aseptic culture is an abnormal condition that likely influences exudation. Studies in soil and field conditions measure the equilibrium concentration resulting from excretion, uptake, and adsorption to the medium. This should not suggest that quantification of plant exudates in soil or similar field conditions is superior or inferior. The two approaches give very different information, both practical in their own ways. Seed exudates are simpler and more consistent, which makes them good model systems for the chemical ecology of young plants. Root exudates, however, describe a much broader portion of the plant's lifecycle, but trades the environmental relevance for the comparative difficulty in study.

Cultivation of legumes is an important component of global agriculture and is unique in its use of nitrogen-fixing rhizobia. Because of this, farmers apply inoculants to seeds and land to take advantage of the symbiosis (51). Human use of microbial inoculants has a long history, starting with a patent of rhizobial culture submitted in 1895 (52). While rhizobia are common in the soil, a field that has not cultivated a particular legume may be bereft of the strains capable of nodulating the plant. Inoculation solves this problem by furnishing an abundance of the proper rhizobia. Inoculation also provides humans with control over which symbionts inoculate their crops. Strains of rhizobia are not equal in their readiness to fix nitrogen. In some cases, the trade between plant and microbe approaches parasitism (53, 54). A major goal of symbiotic nitrogen fixation (SNF) research is to develop strains that efficiently fix nitrogen and can out-compete native strains that would not optimize crop yields (54, 55). Out-competing native strains requires emphasis, and this is where information on chemotaxis becomes critical. As stated before, strains exhibiting chemotaxis will better out-compete their neighbors (56). We have already accumulated a wealth of information about the chemical cues that are the basis for plant-microbe recruitment (2-4, 11).

Additionally, significant progress has been made on the mapping of plant root exudates (38, 39, 50, 57-59). Combining this information enlightens the modification of organisms for the optimization of legume inoculation, and by extension, crop yield. Bacterial MCPs can be modified to alter ligand profiles or attractant preferences to the end of developing ultra-competitive strains that effectively deliver high-efficiency nitrogen-fixing systems (60). This approach can also draw from the abundant work on MCPs from other species that are absent in some rhizobia (18, 61-65). An alternative, though somewhat more dubious and involved method would see crops with modified exudate profiles tailored to their specific endosymbionts. It should still be noted that since the plant supplies the carbon energy source to the bacteroids, the capture of nitrogen from the atmosphere costs energy that might otherwise go to yields (66). However, the efficiency of this conversion can be improved and the loss in yields may be outweighed by the reduction in fertilizer cost (67). Even so, this problem could be further circumvented by a related new area of promise called the plant growth promoting rhizobacteria (PGPR). PGPR are any prokaryotic species that occupy the rhizosphere and enhance the growth and health of the plant. This is achieved through numerous mechanisms including antibiotic production, mineral acquisition through siderophores and solubilization, plant hormone stimulation, stress tolerance (heat, cold, salt, flooding), induced systemic resistance, and crowding out of pathogens (68-72). Unlike rhizobia, PGPR are not limited to a specific host range, although their effects may vary depending on plant type. Specific organisms used include canonical rhizobia like *Azospirillum* and *Bradyrhizobium*, as well as members of genera typically thought of as pathogens, such as *Serratia* spp. and *Burkholderia* spp. A proper discussion of these strains and their mechanisms requires much more space than is available here. However, like specific rhizobial inoculants for legumes, PGPR must survive in the soil and compete for space at the host interface. Chemotaxis and motility assist both objectives. Rhizobacteria must be in relative proximity to their host, and chemotaxis to root or seed exudates allows fine localization to the plant. Survival in the soil away from a host is also an important trait for an effective symbiont. Populations of microbial mutualists can outlast their annual or tilled hosts and survive to colonize the next season's crops. In these scenarios, inoculants can become lasting residents and support crop growth for numerous generations. Some application strategies deposit the inoculants in bulk soil away from the host. In non-hospitable soil types, rhizobial populations can be depleted in as little as 1 - 2 months (73). Nodulating bacteria, however can be

detected years after inoculation (74). In both scenarios, chemotaxis is a greatly beneficial trait for fitness in the soil environments (75-77).

Ultimately, PGPR and other inoculants are important tools in a movement that is set to supplant the green revolution of the 1960s. The emphasis on modern agricultural movements is to do more with less. Improving crop fertilizer uptake reduces production costs and eliminates runoff pollutants. Beneficial microorganisms can be recruited to help fight or ameliorate the effects of pathogens, obviating toxic pesticides. Perhaps in the future, microorganisms will be given credit as stewards of the soil and maintainers of crop productivity.

## REFERENCES

1. Webb BA, Hildreth S, Helm RF, Scharf BE. *Sinorhizobium meliloti* chemoreceptor McpU mediates chemotaxis toward host plant exudates through direct proline sensing. *Applied and environmental microbiology*. 2014;80(11):3404-15.
2. Webb BA, Compton KK, Castañeda Saldaña R, Arapov T, Ray WK, Helm RF, et al. *Sinorhizobium meliloti* chemotaxis to quaternary ammonium compounds is mediated by the chemoreceptor McpX. *Molecular microbiology*. 2017;103(2):333-46.
3. Webb BA, Compton KK, Del Campo JSM, Taylor D, Sobrado P, Scharf BE. *Sinorhizobium meliloti* chemotaxis to multiple amino acids is mediated by the chemoreceptor McpU. *Molecular plant-microbe interactions : MPMI*. 2017;30(10):770-7.
4. Compton KK, Hildreth SB, Helm RF, Scharf BE. *Sinorhizobium meliloti* chemoreceptor mcpv senses short-chain carboxylates via direct binding. *J Bacteriol*. 2018;200(23).
5. Kape R, Parniske M, Werner D. Chemotaxis and nod gene activity of *Bradyrhizobium japonicum* in response to hydroxycinnamic acids and isoflavonoids. *Applied and environmental microbiology*. 1991;57(1):316-9.
6. Dharmatilake AJ, Bauer WD. Chemotaxis of *Rhizobium-meliloti* towards nodulation gene-inducing compounds from alfalfa roots. *Applied and environmental microbiology*. 1992;58(4):1153-8.
7. Caetanoanollés G, Cristestes DK, Bauer WD. Chemotaxis of *Rhizobium-meliloti* to the plant flavone luteolin requires functional nodulation genes. *J Bacteriol*. 1988;170(7):3164-9.
8. Aguilar JMM, Ashby AM, Richards AJM, Loake GJ, Watson MD, Shaw CH. Chemotaxis of *Rhizobium-leguminosarum* biovar phaseoli towards flavonoid inducers of the symbiotic nodulation genes. *J Gen Microbiol*. 1988;134:2741-6.
9. Meier VM, Muschler P, Scharf BE. Functional analysis of nine putative chemoreceptor proteins in *Sinorhizobium meliloti*. *J Bacteriol*. 2007;189(5):1816-26.
10. Meier VM, Scharf BE. Cellular localization of predicted transmembrane and soluble chemoreceptors in *Sinorhizobium meliloti*. *J Bacteriol*. 2009;191(18):5724-33.
11. Webb BA, Helm RF, Scharf BE. Contribution of individual chemoreceptors to *Sinorhizobium meliloti* chemotaxis towards amino acids of host and nonhost seed exudates. *Molecular plant-microbe interactions : MPMI*. 2016;29(No. 3):231-9.
12. Adler J. A method for measuring chemotaxis and use of the method to determine optimum conditions for chemotaxis by *Escherichia coli*. *J Gen Microbiol*. 1973;74(1):77-91.
13. Wiseman T, Williston S, Brandts JF, Lin LN. Rapid measurement of binding constants and heats of binding using a new titration calorimeter. *Anal Biochem*. 1989;179(1):131-7.
14. Matilla MA, Martin-Mora D, Krell T. The use of isothermal titration calorimetry to unravel chemotactic signalling mechanisms. *Environmental Microbiology*. 2020.
15. Garcia V, Reyes-Darias JA, Martin-Mora D, Morel B, Matilla MA, Krell T. Identification of a chemoreceptor for C2 and C3 carboxylic acids. *Applied and environmental microbiology*. 2015;81(16):5449-57.
16. Futrelle RP, Berg HC. Specification of gradients used for studies of chemotaxis. *Nature*. 1972;239(5374):517-8.
17. Reyes-Darias JA, Yang YL, Sourjik V, Krell T. Correlation between signal input and output in PctA and PctB amino acid chemoreceptor of *Pseudomonas aeruginosa*. *Molecular microbiology*. 2015;96(3):513-25.



18. Lacal J, Alfonso C, Liu XX, Parales RE, Morel B, Conejero-Lara F, et al. Identification of a chemoreceptor for tricarboxylic acid cycle intermediates. *J Biol Chem.* 2010;285(30):23124-34.
19. Luu RA, Schomer RA, Brunton CN, Truong R, Ta AP, Tan WA, et al. Hybrid two-component sensors for identification of bacterial chemoreceptor function. *Applied and environmental microbiology.* 2019;85(22).
20. Cieslinski G, Van Rees KCJ, Szmigielska AM, Krishnamurti GSR, Huang PM. Low-molecular-weight organic acids in rhizosphere soils of durum wheat and their effect on cadmium bioaccumulation. *Plant Soil.* 1998;203(1):109-17.
21. Garcia JAL, Barbas C, Probanza A, Barrientos ML, Manero FJG. Low molecular weight organic acids and fatty acids in root exudates of two *Lupinus* cultivars at flowering and fruiting stages. *Phytochem Analysis.* 2001;12(5):305-11.
22. Chiu CY, Wang MK, Hwong JL, King HB. Physical and chemical properties in rhizosphere and bulk soils of *Tsuga* and *Yushania* in a temperate rain forest. *Commun Soil Sci Plan.* 2002;33(11-12):1723-35.
23. Li XL, Chen XM, Liu X, Zhou LC, Yang XQ. Characterization of soil low-molecular-weight organic acids in the Karst rocky desertification region of Guizhou Province, China. *Front Env Sci Eng.* 2012;6(2):195-203.
24. Ortega A, Zhulin IB, Krell T. Sensory repertoire of bacterial chemoreceptors. *Microbiol Mol Biol Rev.* 2017;81(4).
25. Zatakia HM, Arapov TD, Meier VM, Scharf BE. Cellular stoichiometry of methyl-accepting chemotaxis proteins in *Sinorhizobium meliloti*. *J Bacteriol.* 2018;200(6).
26. Alloing G, Travers I, Sagot B, Le Rudulier D, Dupont L. Proline betaine uptake in *Sinorhizobium meliloti*: Characterization of Prb, an opp-like ABC transporter regulated by both proline betaine and salinity stress. *J Bacteriol.* 2006;188(17):6308-17.
27. Hartwig UA, Maxwell CA, Joseph CM, Phillips DA. Chrysoeriol and luteolin released from alfalfa seeds induce nod genes in *Rhizobium-meliloti*. *Plant Physiol.* 1990;92(1):116-22.
28. Hartwig UA, Joseph CM, Phillips DA. Flavonoids released naturally from alfalfa seeds enhance growth-rate of *Rhizobium-meliloti*. *Plant Physiol.* 1991;95(3):797-803.
29. Charles TC, Cai GQ, Aneja P. Megaplasmid and chromosomal loci for the PHB degradation pathway in *Rhizobium (Sinorhizobium) meliloti*. *Genetics.* 1997;146(4):1211-20.
30. Dunn MF. Tricarboxylic acid cycle and anaplerotic enzymes in rhizobia. *Fems Microbiol Rev.* 1998;22(2):105-23.
31. Geddes BA, Oresnik IJ. Physiology, genetics, and biochemistry of carbon metabolism in the alphaproteobacterium *Sinorhizobium meliloti*. *Can J Microbiol.* 2014;60(8):491-507.
32. Georgi CE, Ettinger JM. Utilization of carbohydrates and sugar acids by the rhizobia. *J Bacteriol.* 1941;41(3):323-40.
33. Iyer B, Rajput MS, Jog R, Joshi E, Bharwad K, Rajkumar S. Organic acid mediated repression of sugar utilization in rhizobia. *Microbiol Res.* 2016;192:211-20.
34. Jia GM, Zhang BL, Niu JT, Wang LM, Chen FQ. Soil labile organic carbon fractions in rhizosphere soil in citrus plantations in the Three Gorges Reservoir Area. *Agroforest Syst.* 2015;89(6):1097-105.
35. Jones DL, Nguyen C, Finlay RD. Carbon flow in the rhizosphere: carbon trading at the soil-root interface. *Plant Soil.* 2009;321:5-33.
36. R. P, D. P. The sites of excretion of ninhydrin-positive substances by broad bean seedlings. *Plant Soil.* 1961;13:391-6.

37. Czarnota MA, Rimando AM, Weston LA. Evaluation of root exudates of seven sorghum accessions. *J Chem Ecol.* 2003;29(9):2073-83.
38. Badri DV, Vivanco JM. Regulation and function of root exudates. *Plant Cell Environ.* 2009;32(6):666-81.
39. Canarini A, Kaiser C, Merchant A, Richter A, Wanek W. Root exudation of primary metabolites: mechanisms and their roles in plant responses to environmental stimuli. *Front Plant Sci.* 2019;10:157.
40. Hawes MC, Gunawardena U, Miyasaka S, Zhao X. The role of root border cells in plant defense. *Trends Plant Sci.* 2000;5(3):128-33.
41. Vicre M, Santaella C, Blanchet S, Gateau A, Driouich A. Root border-like cells of *Arabidopsis*. Microscopical characterization and role in the interaction with rhizobacteria. *Plant Physiol.* 2005;138(2):998-1008.
42. O'Neal L, Vo L, Alexandre G. Specific root exudate compounds sensed by dedicated chemoreceptors shape *Azospirillum brasilense* chemotaxis in the rhizosphere. *Appl Environ Microbiol.* 2020;86(15).
43. Götz R, Limmer N, Ober K, Schmitt R. Motility and chemotaxis in 2 strains of rhizobium with complex flagella. *J Gen Microbiol.* 1982;128(Apr):789-98.
44. Radchuk V, Borisjuk L. Physical, metabolic and developmental functions of the seed coat. *Front Plant Sci.* 2014;5:510.
45. Maxwell CA, Phillips DA. Concurrent synthesis and release of nod-gene-inducing flavonoids from alfalfa roots. *Plant Physiol.* 1990;93(4):1552-8.
46. Shen H, He LF, Sasaki T, Yamamoto Y, Zheng SJ, Ligaba A, et al. Citrate secretion coupled with the modulation of soybean root tip under aluminum stress. Up-regulation of transcription, translation, and threonine-oriented phosphorylation of plasma membrane H<sup>+</sup>-ATPase. *Plant Physiol.* 2005;138(1):287-96.
47. Ryan PR, Raman H, Gupta S, Horst WJ, Delhaize E. A second mechanism for aluminum resistance in wheat relies on the constitutive efflux of citrate from roots. *Plant Physiol.* 2009;149(1):340-51.
48. Liu MY, Lou HQ, Chen WW, Pineros MA, Xu JM, Fan W, et al. Two citrate transporters coordinately regulate citrate secretion from rice bean root tip under aluminum stress. *Plant Cell Environ.* 2018;41(4):809-22.
49. Phillips DA, Fox TC, King MD, Bhuvaneshwari TV, Teuber LR. Microbial products trigger amino acid exudation from plant roots. *Plant Physiol.* 2004;136(1):2887-94.
50. Carvalhais LC, Dennis PG, Fedoseyenko D, Hajirezaei MR, Borriss R, von Wiren N. Root exudation of sugars, amino acids, and organic acids by maize as affected by nitrogen, phosphorus, potassium, and iron deficiency. *J Plant Nutr Soil Sc.* 2011;174(1):3-11.
51. Catroux G, Hartmann A, Revellin C. Agronomic benefits of rhizobial inoculant use over nitrogen fertilizer application in tropical soybean. *Plant Soil.* 2001;230:21-30.
52. Nobbe F, Hiltner L, inventors. Inoculation of the soil for the cultivation of leguminous plants. United States of America 1895.
53. Denison RF, Kiers ET. Lifestyle alternatives for rhizobia: mutualism, parasitism, and forgoing symbiosis. *FEMS Microbiol Lett.* 2004;237(2):187-93.
54. Irisarri P, Cardozo G, Tartaglia C, Reyno R, Gutierrez P, Lattanzi FA, et al. Selection of competitive and efficient rhizobia strains for white clover. *Front Microbiol.* 2019;10:768.

55. Mus F, Crook MB, Garcia K, Garcia Costas A, Geddes BA, Kouri ED, et al. Symbiotic nitrogen fixation and the challenges to its extension to nonlegumes. *Appl Environ Microbiol.* 2016;82(13):3698-710.
56. Ames P, Bergman K. Competitive advantage provided by bacterial motility in the formation of nodules by *Rhizobium-meliloti*. *J Bacteriol.* 1981;148(2):728-9.
57. Fisher RF, Long SR. Rhizobium - Plant Signal Exchange. *Nature.* 1992;357(6380):655-60.
58. Kong CH, Zhang SZ, Li YH, Xia ZC, Yang XF, Meiners SJ, et al. Plant neighbor detection and allelochemical response are driven by root-secreted signaling chemicals. *Nat Commun.* 2018;9(1):3867.
59. Nelson EB. Microbial dynamics and interactions in the spermosphere. *Annual review of phytopathology.* 2004;42:271-309.
60. Bi S, Lai L. Bacterial chemoreceptors and chemoeffectors. *Cell Mol Life Sci.* 2015;72(4):691-708.
61. Martin-Mora D, Reyes-Darias JA, Ortega A, Corral-Lugo A, Matilla MA, Krell T. McpQ is a specific citrate chemoreceptor that responds preferentially to citrate/metal ion complexes. *Environ Microbiol.* 2016;18(10):3284-95.
62. Fernandez M, Morel B, Corral-Lugo A, Krell T. Identification of a chemoreceptor that specifically mediates chemotaxis toward metabolizable purine derivatives. *Mol Microbiol.* 2016;99(1):34-42.
63. Ortega Á, Zhulin IB, Krell T. Sensory repertoire of bacterial chemoreceptors. *Microbiology and Molecular Biology Reviews.* 2017;81(4).
64. Gavira JA, Ortega A, Martin-Mora D, Conejero-Muriel MT, Corral-Lugo A, Morel B, et al. Structural basis for polyamine binding at the dcache domain of the mcpu chemoreceptor from *Pseudomonas putida*. *J Mol Biol.* 2018;430(13):1950-63.
65. Ni B, Huang Z, Fan Z, Jiang CY, Liu SJ. *Comamonas testosteroni* uses a chemoreceptor for tricarboxylic acid cycle intermediates to trigger chemotactic responses towards aromatic compounds. *Mol Microbiol.* 2013;90(4):813-23.
66. Udvardi M, Poole PS. Transport and metabolism in legume-rhizobia symbioses. *Annu Rev Plant Biol.* 2013;64:781-805.
67. Pinochet X, Arnaud F, Cleyetmarel JC. Competition for nodule occupancy of introduced *Bradyrhizobium-japonicum* strain smgs1 in french soils already containing *Bradyrhizobium-japonicum* strain G49. *Can J Microbiol.* 1993;39(11):1022-8.
68. Schwartz AR, Ortiz I, Maymon M, Herbold CW, Fujishige NA, Vijanderan JA, et al. *Bacillus simplex*-a little known pgpb with anti-fungal activity-alters pea legume root architecture and nodule morphology when coinoculated with rhizobium leguminosarum bv.viciae. *Agronomy-Basel.* 2013;3(4):595-620.
69. Maymon M, Martinez-Hidalgo P, Tran SS, Ice T, Craemer K, Anbarchian T, et al. Mining the phytomicrobiome to understand how bacterial coinoculations enhance plant growth. *Front Plant Sci.* 2015;6:784.
70. Drogue B, Dore H, Borland S, Wisniewski-Dye F, Prigent-Combaret C. Which specificity in cooperation between phytostimulating rhizobacteria and plants? *Research in microbiology.* 2012;163(8):500-10.
71. Backer R, Rokem JS, Ilangumaran G, Lamont J, Praslickova D, Ricci E, et al. Plant growth-promoting rhizobacteria: context, mechanisms of action, and roadmap to commercialization of biostimulants for sustainable agriculture. *Frontiers in Plant Science.* 2018;9.

72. Gouda S, Kerry RG, Das G, Paramithiotis S, Shin HS, Patra JK. Revitalization of plant growth promoting rhizobacteria for sustainable development in agriculture. *Microbiol Res.* 2018;206:131-40.
73. Da HN, Deng SP. Survival and persistence of genetically modified *Sinorhizobium meliloti* in soil. *Appl Soil Ecol.* 2003;22(1):1-14.
74. Hirsch PR, Spokes JD. Survival and dispersion of genetically modified rhizobia in the field and genetic interactions with native strains. *FEMS Microbiology Ecology.* 1994;15(1-2):147-59.
75. Yang L, Chen X, Zeng X, Radosevich M, Ripp S, Zhuang J, et al. Surface-adsorbed contaminants mediate the importance of chemotaxis and haptotaxis for bacterial transport through soils. *Front Microbiol.* 2019;10:2691.
76. Poole P, Ramachandran V, Terpolilli J. Rhizobia: from saprophytes to endosymbionts. *Nat Rev Microbiol.* 2018;16(5):291-303.
77. Guo XQ, Chen K, Wen Y, Li R, Li SP, Jiang JD. Comparison of in-situ biodegrading abilities of *Pseudomonas putida* mutants: *leuB(-)* auxotroph, *fliC(-)* non-motility, and *cheA(-)* non-chemotaxis. *Int Biodeter Biodegr.* 2009;63(5):576-81.